Climate policies between carbon prices, oil rents and urban dynamics

H. Waisman

Thèse présentée pour obtenir le grade de Docteur
de l'Ecole des Hautes Etudes en Sciences Sociales

Discipline : Economie

Henri WAISMAN

Les politiques climatiques entre prix du carbone, rentes pétrolières et dynamiques urbaines

Climate policies between carbon prices, oil rents and urban dynamics

----

Thèse dirigée par Jean-Charles Hourcade, avec le co-encadrement de Fabio Grazi

réalisée au CIRED, 45bis avenue de la Belle Gabrielle, 94736 Nogent-sur-Marne

présentée et soutenue publiquement le 17 avril 2012 devant le jury composé de :

M. Patrick CRIQUI, Directeur de Recherche CNRS — *Rapporteur*

M. Jean LATERRASSE, Professeur à l'université Paris Est. — *Rapporteur*

M. John WEYANT, Professeur à l'université de Stanford, USA — *Examinateur*

M. Shobhakar DHAKAL, Directeur exécutif du Global Carbon Project,
National Institute for Environmental Studies, Japon — *Examinateur*

M. Jean-Charles HOURCADE, Directeur de Recherche CNRS et
Professeur à l'EHESS — *Directeur de thèse*

En pensant à Raphaël, Eva et tous les autres…

Remerciements

Avant toute chose, je remercie Patrick Criqui et Jean Laterrasse d'avoir accepté d'être rapporteurs de mon travail, tâche rendue d'autant plus ardue que des délais très courts leur ont été imposés pour la lecture critique du manuscrit. Merci à eux de s'être conformés à ce principe fondamental du CIRED selon lequel seule l'urgence permet un travail abouti.

Thanks also to the two members of the jury, John Weyant and Shobhakar Dhakal, who have accepted to come to Paris to attend this PhD defense from their land of Far-Far Away (USA and Japan, respectively). Both have welcomed me very nicely in the scientific community and have largely contributed to the work presented here through very interesting discussions about our common scientific interests

Cette thèse n'aurait évidemment jamais vu le jour sans Jean-Charles Hourcade, directeur de thèse et directeur de laboratoire iconoclaste s'il en est, qui a pour habitude d'accepter les nouveaux venus sur la base de son intuition (rarement mise en défaut) qu'ils trouveront leur place dans la grande famille du CIRED. Dans mon cas, cette première rencontre, un entretien d'une demi-heure auquel je n'ai absolument rien compris sur le moment, contenait déjà toutes les idées qui se sont concrétisées dans cette thèse (et notamment l'importance du « prix du mètre carré » !) et, depuis lors, ma principale tâche a été de construire un dictionnaire me permettant de traduire les digressions hourcadiennes en directions de recherche. Mais Jean-Charles Hourcade n'est pas seulement un directeur de thèse, c'est aussi un puits de savoir avec qui les discussions ont du mal à ne pas dériver sur des sujets aussi cruciaux que la religion, la politique, la famille ou l'importance des spécificités régionales. Tous ceux qui l'ont fréquenté vous le diront, c'est pour cette dimension humaine spécifique qu'on l'adore grâce à la pertinence et la richesse de ses analyses, ou qu'on le déteste (surtout le vendredi soir à 21h).

Cette thèse n'aurait jamais pu arriver à son terme sans l'aiguillon instillé en permanence et la contribution concrète de Fabio Grazi. Il a plus que largement contribué à l'écriture de la deuxième partie de cette thèse, mais il a aussi relu en détail l'intégralité de la thèse pour m'aider à améliorer la clarté de mes explications dans le fond et dans la forme, ligne par ligne, et a su me transmettre son opiniâtreté et à sa confiance inconditionnelles et communicatives.

Mais Fabio a été bien plus qu'un collègue de travail, il m'a aussi beaucoup appris sur les relations humaines et a partagé certaines des remises en question les plus fondamentales au cours de cette tranche de vie qu'est la thèse. Son principal défaut, que je ne peux évacuer ici, est son manque de discernement qui le conduit à soutenir l'équipe d'Italie de football et le Milan AC, deux erreurs dramatiques qui ont été une des plus importantes sources de division entre nous

Si le travail de thèse est traditionnellement considéré comme un travail personnel, c'est beaucoup moins vrai dans mon cas, puisque rien n'aurait été possible sans les différents contributeurs du modèle IMACLIM, au cœur des analyses qui y sont menées. En premier lieu, j'ai une pensée sincèrement émue pour Renaud Crassous et Olivier Sassi, les patriarches d'IMACLIM, qui, au moment où je suis arrivé au CIRED, ont posé les fondements du monstre sur lequel nous sommes encore en train de travailler aujourd'hui, sur la base d'intuitions riches et avec une volonté et une foi qui forcent le respect. Ils m'ont accueilli chaleureusement au sein d'une équipe IMACLIM qui ne comptait à l'époque que deux membres mais qui s'est depuis fortement élargie. Parmi ces membres de la glorieuse aventure IMACLIM, je voudrais ici remercier Christophe Cassen qui assume la lourde tâche de coordonner les efforts disparates de développement et permet, par son volontarisme, de maintenir une vraie cohésion et une vraie cohérence au travail d'équipe. Un grand merci à tous les « ouvriers » d'IMACLIM, dont je suis fier de faire partie et qui ont grandement contribué à cette thèse grâce aux discussions et au travail en commun enrichissants qu'ils ont permis : Ruben Bibas, Gaëtan Giraudet, Céline Guivarch, Julien Lefèvre, Aurélie Méjean, Elsa Mosseri, Julie Rozenberg et Adrien Vogt-Schilb. Une pensée spéciale pour Cédric Allio, avec qui la collaboration quotidienne se révèle très agréable et enrichissante, et qui a beaucoup contribué aux différents travaux menés sur les questions urbaines en faisant preuve d'une grande maturité et dont les prises d'initiatives ont permis des avancées cruciales dans les travaux en cours.

De nombreuses personnes au CIRED ont également joué plus indirectement un rôle important dans cette thèse : Philippe Quirion et Sandrine Mathy qui ont facilité mon acclimatation en m'accueillant dans leur « bureau du fond », forteresse sacrée du CIRED, où la recherche académique et l'engagement associatif se combinent avec bonheur pour donner un sens pragmatique à nos travaux ; Paolo Avner, qui a accepté avec grande patience de relire, re-relire et corriger mon anglais approximatif pour le rendre acceptable; Valentin Przyluski, avec qui les discussions politiques animées, notamment sur la démondialisation, ont enrichi ma

compréhension de ce monde mystérieux qu'est l'engagement militant et nourri la réflexion sur la façon dont notre travail peut être mis en application; Yaël Serfaty, qui a largement contribué par sa sérénité excitée à partager et affronter les moments d'irrationalité assumée du travail au CIRED et qui m'a démontré à quel point l'Afrique du Nord peut être au final si proche de l'Europe de l'Est; Eléonore Tyma, qui constitue un rouage essentiel aux travaux du CIRED en supportant avec calme les demandes récurrentes d'envoi de courrier, de paquets de café ou de stylos divers ; Catherine Boemare avec qui j'ai pu partager les vicissitudes de la collaboration étroite avec le Grand Chef ; Jean-Pierre Mailles, pilier de la SMASH, qui a toujours fait preuve de flexibilité pour faciliter la vie des chercheurs du CIRED ; Tarik Tazdaït et Franck Nadaud, grâce à qui j'ai pu mieux appréhender et décrypter le grand complot mondial mené par les puissants contre les pauvres hères que nous sommes ; enfin tous les partenaires de pause café et de déjeuners avec qui les discussions enrichissantes ont servi à nourrir les réflexions sur le travail de recherche mais aussi sur des questions aussi cruciales que l'amitié relative que l'on peut éprouver pour un saumon ou un poulet.

Mais le CIRED n'est pas seulement un lieu de travail, c'est aussi un endroit magique où j'ai pu trouver de vrais amis qui m'ont accompagné tout au long de cette période, ont partagé les bons moments mais aussi les périodes plus difficiles en m'accordant toujours leur soutien et en m'aidant à avancer dans la réflexion sur le monde dans lequel nous évoluons.

Je garde ainsi un souvenir ému des pauses café interminables de mes premières années au CIRED où j'ai pu apprendre, au contact de Damien et Guy, la valeur du regard critique sur les principes dont on est intimement convaincu et l'intérêt des discussions contradictoires pour aboutir à un raisonnement cohérent. Bien que s'étant égaré depuis dans le soutien au PSG (personne n'est parfait), Damien reste ainsi un ami fidèle, dont l'amitié constante et valorisante a constitué un moteur essentiel des aspects positifs de mon évolution, et avec qui la complicité sur les valeurs essentielles s'est imposée très rapidement sans jamais être démentie. Quant à Guy, sa posture de questionnement critique et son amour immodéré de la discussion resteront pour moi un modèle à suivre dans l'optique de progresser sur la compréhension du monde qui nous entoure. Un grand merci également à Elsa, grâce à qui j'ai pu découvrir les joies de Porquerolles et qui m'a continuellement encouragé, à sa façon, dans l'avancée de cette thèse (à raison d'une vanne par soirée, pas plus…).

Plus récemment, Manu a partagé les difficultés et les espoirs associés à la progression finale de la thèse en me faisant bénéficier de sa force de caractère, de son perfectionnisme et de sa capacité à affronter les difficultés qui surgissent sans prévenir tout en menant de front une carrière académique et un destin de rock-star, un vrai exemple !

Mais, la thèse n'est pas seulement un travail, c'est aussi et surtout une vraie tranche de vie au cours de laquelle les amis de toujours ont parfaitement joué leur rôle. Tout d'abord, Lolo, MC, Louisa, Sara, Numa, Ben, Scaloup, amis depuis le lycée, ont accompagné la progression de cette thèse. Une pensée spéciale pour Delphine qui a contribué de façon décisive à mon intérêt pour les questions environnementales, m'a encouragé quand il le fallait à prendre le risque de me lancer dans une reconversion, incertaine à l'époque, de la physique vers l'économie de l'environnement et a indéniablement joué, de près puis de loin, un rôle déterminant dans les aspects positifs de mon évolution.

Laurent a également joué un rôle prépondérant au cours de cette thèse au travers des nombreuses soirées où je me suis senti chez lui comme chez moi pour partager des repas, des séries et des discussions innombrables, et en me faisant profiter de son dynamisme et de sa « chance légendaire » pour passer d'innombrables bons moments ensemble. Une pensée également pour Guillaume avec qui les discussions prennent immanquablement un tour polémique très stimulant, et pour Gab dont l'humanité et la fragilité sont un véritable moteur.

Un grand merci à celles et ceux qui ont égayé cette période parfois délicate en me faisant profiter de leur dynamisme et de leur amitié, en jouant un rôle apaisant dans les périodes de doute et stimulant dans les bonnes périodes. C'est le cas d'Elsa, qui prouve à elle seule l'importance des conférences internationales et des projets du ministère, et qui m'a toujours encouragé et valorisé en jouant indéniablement un rôle apaisant de par sa force de caractère et son dynamisme ; Mélanie, qui a aussi été une source d'inspiration en opérant un mélange subtil entre indépendance et fragilité, et m'a permis de connaître les Enfants de la Goutte D'Or, où j'ai pu rencontrer des gens motivés par le bien commun, Lydie, Wardine, Bénédicte et tous les enfants, dont la volonté sans faille ont indéniablement joué un rôle essentiel dans la foi dans le bien collectif qui préside à cette thèse ; enfin, merci aux Jeunes du Stade, et particulièrement à Pedro et David, qui donnent un vrai sens à l'engagement associatif et collectif sans lequel la performance individuelle perd son sens.

Il existe également des personnes qui échappent à toute catégorie comme Mériem, membre éminente du CIRED, mais bien plus que cela. Aucun mot ne peut décrire son importance dans le cours de cette thèse, que ce soit au travers de discussions sur des sujets scientifiques ou d'échanges sur des problématiques plus fondamentales touchant à la vie personnelle et à sa perception subjective. Sans elle, sans les discussions jusqu'à pas d'heure sur IMACLIM et le sens de la vie, sans son soutien et son affection désintéressés, sans son amitié inconditionnelle dans les périodes les plus délicates, nul doute que cette thèse ne serait jamais arrivée à son terme. C'est rarissime de trouver quelqu'un qui vous comprend parfaitement, qui se préoccupe de vous et qui vous aide vraiment, moi j'ai eu cette chance avec Mériem!

Enfin, cette thèse n'aurait jamais vu le jour sans le soutien inconditionnel de ma famille. Mes parents, qui n'ont cessé de m'encourager et de me valoriser depuis ma plus tendre enfance, ont réussi le subtil mélange d'affection et de tendresse me permettant de m'épanouir dans les périodes fastes tout en me sentant soutenu dans les périodes les plus délicates. Mon frère Jonathan et ma belle-sœur Taryna ont également joué un rôle stabilisateur essentiel en me témoignant leur affection permanente et en montrant la voie d'une évolution épanouie. Enfin ma sœur Julie, mon portrait « craché » et pourtant complètement différente, a été une source constante d'inspiration notamment grâce à sa capacité à évoluer positivement malgré les difficultés.

Pour finir, un grand merci à tous les membres de la nouvelle génération, tous plus extraordinaires les uns que les autres, qui donnent une vraie signification à ces travaux sur le devenir de notre société pour les générations futures : Raphaël, Eva, Nassim, Rebecca, Maïa, Gabriel, Alix...

Abstract

This thesis investigates the effects of constraints imposed on economic interactions by limitations due to natural resources, among which oil and urban land play a curcial role in the context of climate change. These dimensions, often neglected in existing analyses, have an ambiguous effect since they suggest both the risk of enhanced costs if carbon limitations reinforce the sub-optimalities caused by pre-existing constraints, but also, conversely, the possibility of co-benefits if the climate policy helps to correct some pre-existing imperfections of socio-economic trajectories. To investigate this issue, an innovative modeling framework of the enrgy-economy interactions is elaborated that embarks the specificities of the deployment of oil production capacities and the issues related to the spatial organization in urban areas. We demonstrate that, beyond the carbon price, the costs of climate policy essentially depend on the sequencing of complementary measures, with a crucial role of spatial policy designed to control transport-related emissions through mobility.

Résumé

Cette thèse analyse les effets de contraintes sur les interactions économiques imposées par les ressources naturelles, parmi lesquelles le pétrole et la terre urbaine jouent un rôle crucial dans le contexte du changement climatique. Ces dimensions, souvent négligées dans les études existantes, ont un effet ambigu puisqu'elles suggèrent à la fois le risque de coûts exacerbés par les contraintes, mais aussi de potentiels de co-bénéfices si la politique climatique aide à corriger certaines sous-optimalités des trajectoires socio-économiques. Pour analyser ces effets, une architecture de modélisation innovante des interactions énergie-économie est développée, qui prend en compte les spécificités du déploiement des capacités de production pétrolières et les enjeux de l'organisation spatiale dans les aires urbaines. Nous montrons en particulier que, au-delà de la tarification du carbone, les coûts d'une politique climatique dépendent du tuilage des différentes mesures d'accompagnement mises en œuvre, avec un rôle essentiel pour les politiques spatiales construites pour contrôler les émissions liées au transport *via* la mobilité

List of Tables

List of Figures

General introduction

Twenty years after the adoption of the United Nations Framework Convention on Climate Change (UNFCCC) at the 1992 Earth Summit in Rio, its ojectives are more topical than ever, namely "achieve […] stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system." Indeed, progress in the scientific understanding of climate change and of its impacts has permitted to reach a large consensus to acknowledge both the responsibility of emissions from human activity (essentially from fossil fuel burning) in the acceleration of climate change over the last decades[1] and its potentially catastrophic socio-economic effects, especially in developing countries[2]. A review of the literature suggests in addition that the socio-economic costs of carbon reduction measures may remain limited to a few percentage points of GDP even with the most ambitious stabilization objectives[3].

But, in parallel, the annual Conferences of the Parties, created by UNFCCC to organize the political negotiations on climate issues, have failed in their mission to elaborate a tangible agreement on emission reduction objectives and policies to be implemented for this purpose. The more recent occurences in Copenhague, Cancun and Durban have even reinforced the scepticism about this political process since, despite a wide mobilization and numerous declarations of intent, no significant political advancement could be made. Here lies the paradox of climate issues: the scientific alerts on the necessity and urgency of acting before the end of the opportunity window for ambitious climate action seem not audible for politicians. We take the point that this paradox is not only a matter of free-riding attitudes, but

[1] The analysis of global carbon cycle proves that man-made fluxes are the main positive contributors to the accumulation of carbon in the atmosphere (AR4, WGI, Table 7.3). This accumulation is estimated to be responsible of 1.6 $W.m^{-2}$ of total forcing of climate over 1750- 2005, to which natural solar irradiance adds noly 0.18 $W.m^{-2}$ (AR4, WGI, Table 2.12). This dominance of human emissions in climate effects is confirmed by climate models, which can reproduce observed trends of global temperature changes only if they use anthropogenic forcings (AR4, WGI,fig SPM4)

[2] A review of the literature demonstrates that impacts of climate change may lead to 1.5% to 3.5% GDP losses at the world level, and that developing countries are the more exposed regions (Section 20.6 in (Yohe et al, 2007)). By including non-market impacts on health and the environment, the possibility that the climate system may be more responsive to greenhouse gas emissions than previously thought and the disproportionate share of the climate-change burden falling on poor regions, Stern (2007) obtains that welfare reductions (measured in equivalent consumption per head) may reach 20%.

[3] Average aggregate losses remain below 5.5% in 2050 (Table SPM-6 in (IPCC, 2007)). These estimates of the economic effects of greenhouse gas emission reductions are based on the use of numerical modeling tools, which are appropriate to provide quantitative assessments of mechanisms at play in complex systems, like the ones considered in energy-economy-environment interactions. This type of analysis has developed in parallel with exponential improvements of computer techniques allowing for a fast growth of the number of models and scenarios, as illustrated by the 177 scenarios compiled in (IPCC, 2007). This multiplicity of results provides a wide spectrum of assumptions and then helps delineate the ranges of uncertainties of the results, which is crucial information in such a controversial issue as the long-term impacts of climate policies.

reveals also a gap between scientific methodologies adopted to assess carbon mitigation costs and politicians' expectations in terms of policy-relevant messages.

This gap is not surprinzing when considering the assumptions of most models used for climate policy assessment:

> "Most models use a global least cost approach to mitigation portfolios and with universal emissions trading, assuming transparent markets, no transaction cost, and thus perfect implementation of mitigation measures throughout the 21st century" (IPCC, 2007, Box SPM.3).

Athough these assumptions can be acceptable to provide a normative vision, they fail to capture the role of constraints which may push economic trajectories away from their first-best trajectory. Among these constraints, we distinguish inertia effects limiting the flexibility of economic adjustments (missing mechanisms) and market imperfections pushing prices above marginal cost (missing market). These two characteristics are particularly important when scarce resources are used as production factor, because they introduce limitations on supply and potentials for market power behaviors according to the regional distribution of the resource.

This thesis considers two scarce resources, oil and urban land use, which have in common a close interplay with transport trends, a crucial sector for climate policy. Indeed, oil markets determine the cost of fuels and the profitability of substitutes, whereas urban forms drive constrained mobility at the local scale. To provide audible insights on the interplay between oil, urban land use and climate policy, we propose methodological advancements to represent that climate policies apply in a world with pre-existing constraints due to:

(a) inertia mechanisms because, respectively, of geopolitical, technical and geological constraints on the time-profile of oil markets, and of spatial and infrastuctural constraints on production relocation and changes in urban structure.

(b) imperfect markets because the owners of a scarce resource can impose rent prices over marginal cost for the exhaustible oil resource and the land close to urban centers, respectively.

These dimensions, often neglected in existing analyses, offer some new perspectives in terms of climate policy analysis. They suggest both the risk of enhanced costs if carbon limitations reinforce the sub-optimalities caused by pre-existing constraints, but also, conversely, the possibility of co-benefits if the climate policy helps to correct some pre-existing imperfections.

Section 1 of this introduction identifies the constraints in interaction driving inertia mechanisms that will be considered in this thesis, and section 2 makes a detour *via* the theory of rents to identify the determinants of pricing mechanisms on imperfect markets. Section 3 synthetizes this information and gives the outline of the thesis.

I. Climate policy and inertia constraints

As we have seen, most models use first-best assumptions and hence do not represent short-term inertia constraints, but much attention has been devoted to the long-run inertia effects related to *technological constraints*. This dimension of climate policies has been investigated in number of modeling exercises, including with hybrid models able to incorporate sector-based expertise on technolgies and technological change into a macroeconomic framework (Hourcade, 2006).[4] More recently, second-best policies have been envisaged to account for *climate negotiation constraints*, which could lead to an agreement involving exemptions or delayed participation for some regions and/or limitations in the use of some low-carbon technologies.[5]

We go one step further by considering the inertias that affect the functioning of economic interactions in the short-run and impose departures from steady growth pathways. More precisely, this means that we consider the interplay between three types of constraints (a) the technical and political constraints, which affects the flexibility of capital adjustements especially under imperfect foresight and impose distortions of economic interactions to satisfy some public objectives; (b) the climate constraint, which has no reason to be optimal with respect to economic dynamics but is rather political decisions imposing a time path for

[4] Among the more recent studies on the topic: the EMF-19 project, which deals with the role of cost and performance of current and future technologies for global climate policies (Weyant, 2004); the IMCP project which studies the role of endogenous technological change on the cost of climate policies (Edenhofer et al, 2006); or the EMF-21 project, which incorporates non-CO2 gases, such as those from land uses and agriculture (Weyant et al. 2007).

[5] This issue is in particular investigated in: the EMF-22 project, which studies the impact of the architecture of climate policies on the possibility to reach a given stabilization target and on the associated costs (Clarke et al, 2010); the ADAM project, which analyses the feasibility and costs of very low stabilization in function of technology availability (Edenhofer et al, 2010); and the RECIPE project, which investigates the effects of delayed participation by certain regions (Jakob et al, 2010), restrictions on the availability of a large set of low-carbon technologies (Tavoni et al., 2011) and quota allocation rules (Luderer et al, 2010).

limitations of CO_2 emissions emitted by human activity; (c) resource constraints, which impose a limit on a scarce resource use as a production factor. We consider more specifically on oil and of urban land, which play a crucial role in climate context.

1- Constraints on the functioning of economic interactions: the second-best economy

When considering the implementation of ambitious climate policies implying important departures from current socio-economic trends, complex interactions in the economic system and unexpected shocks driving the long-term dynamics of prices, quantities and investment decisions cannot but be imperfectly anticipated with the information at the agents' disposal. In this context, it therefore appears appropriate to go beyond the standard assumption of perfect foresight to adopt instead adaptive expectations, which are formed on the extrapolation of past and current trends and are refined as agents get information (for example, on the nature of climate constraints).

These imperfections in expectations play an even more important role when taking into account the constraints that limit the flexibility of economic adjustments and hence prevent the immediate correction of past decisions once information arrives. These constraints concern notably (i) the rigidities on labor markets, which limit the adjustments of labor costs in the production process, (ii) the pace of diffusion of new technologies as constrained by limited R&D potentials and the cumulative effects of learning-by-doing processes, (iii) the renewal of installed technologies, equipments and infrastructure, especially for long-lived equipments in housing and transport sectors, (iv) the behavioral inertias, which limit the pace of structural changes in consumption and production patterns, and (v) the basic needs on vital items (e.g., food, housing), which limit consumption adjustments by imposing a floor level for these specific goods.

Independently from climate concerns, numbers of policies are active and influence socio-economic interactions. They are not necessarily adopted for economic efficiency reasons but rather to satisfy a specific objective, like the supplying of public service that must be satisfied at any cost (e.g, State's commitment to ensure a certain level of public services like security, justice, education, health), the preservation of a certain level of intra- and inter-generational redistribution or the internalization of some indirect economic, social or environmental effects that would otherwise be ignored by economic tradeoffs. These policies move economic trajectories away from their optimal trajectory and distort the structure of prices and quantities (quotas) and/or the technical standards (norms).

In this context, measures adopted for climate concerns have ambiguous consequences, since they can have synergies, or, on the contrary, reinforce pre-existing suboptimalities. This issue is at the core of the debate on the double dividend of climate policies, which is built on the idea that a carbon fiscal reform can potentially lead to absolute economic gains if the redistribution of the product of carbon taxation is used to suppress the more distortive effects of the pre-existing fiscal system (Goulder, 1994; Ligthart, 1998).

The investigation of the interplaying effect of these inertias causing limits in the flexibility of technical adjustments with a climate policy are at the core motivation for a series of analyses carried out at the Center for the Environment and Development (CIRED):

- (Crassous, 2008) demonstrates how those features condition the time profile of carbon prices and of associated economic losses. This study demonstrates in particular that, contrary to most analyzes predicting a steady increase of carbon prices, the second-best nature of the economy leads to: a fast increase in the early period to overcome the constraints imposed by inertias and correct the imperfect expectations about the carbon constraint; a medium-term stagnation permitted by the co-benefits of the climate policy, which contributes to correct some sub-optimalities of the baseline scenarios; and a fast long-term increase to reach the high-cost mitigation potentials in the transport sector. This study also investigates the potentials offered by sectoral and regional differentiation of the climate measures, beyond the standard framework of a unique carbon price.
- (Sassi, 2008) analyzes the interplay between technological evolutions and structural change resulting from endogenous interactions between technical potentials, consumption patterns and location choices. In the context of a climate policy, ambiguous effects appear, since technical change towards less carbon-intensive patterns favors the reduction of emissions, but also negatively affects technical progress in non-energy sectors *via* a crowding-out effect on available investments. these effects are even more important when endogenous technical change increase the risk of technological, structural and behavioral lock-in effects.
- (Guivarch, 2010) considers more specifically the role of rigidities on labor markets and of pre-existing fiscal systems, which prove to be crucial determinants of the

consequences of increases in the cost of energy, like those happening under climate policy. On the one hand, a flexible labor market favors the adjustments of production processes *via* decreases of labor costs, which moderates the surge of production costs; on the contrary, rigidities on labor markets prevent such adjustments and enhance the costs. On the other hand, synergies between climate policies and pre-existing distortions of the fiscal system are investigated on the example of the power sector in India. This sector is characterized by pre-existing subsidies to foster investments and insufficient capacities, which can be corrected at the occasion of a climate policy imposing tariff reforms, demand-side management and improvement in the production/distribution structure.

- (Combet et al, 2010) investigates the interactions between the design of a carbon tax reform, economic efficiency and distributional issues. This study analyzes to what extent a restructuration of the fiscal system on the occasion of a carbon tax reform can have positive effects by redirecting the levy from production to those revenues that are disconnected from production (rent and transfer incomes).

2- The carbon constraint

The choice of a quantified objective for carbon emission reductions and the modalities of its implementation (in terms of regional and sectoral distribution of the efforts) is the consequence of a political appraisal of number of dimensions beyond pure economic efficiency, like energy security, inter-generational equity, a "common but differentiated responsibility" between regions or social aspects (political acceptability of the measures, impacts on development in emerging economies or distributional issues). The importance of these different aspects means that climate policy architecture has no reason to be economically optimal, even more since the uncertainties on climate damages and mitigation costs make it difficult to define *a priori* an optimal climate policy. We adopt a more realistic definition of the climate constraint, as imposed by a political choice on the ultimate stabilization objective. Table 0.1, taken from (IPCC, 2007), describes how the choice of a climate objective translates into a constraint on the time profile of carbon emissions.

Table 0.1: *Characteristics of post-TAR stabilization scenarios*

Category	Radiative forcing (W/m^2)	CO_2 concentration[c)] (ppm)	CO_2-eq concentration[c)] (ppm)	Global mean temperature increase above pre-industrial at equilibrium, using "best estimate" climate sensitivity[b), c)] (°C)	Peaking year for CO_2 emissions[d)]	Change in global CO_2 emissions in 2050 (% of 2000 emissions)[d)]	No. of assessed scenarios
I	2.5-3.0	350-400	445-490	2.0-2.4	2000-2015	-85 to -50	6
II	3.0-3.5	400-440	490-535	2.4-2.8	2000-2020	-60 to -30	18
III	3.5-4.0	440-485	535-590	2.8-3.2	2010-2030	-30 to +5	21
IV	4.0-5.0	485-570	590-710	3.2-4.0	2020-2060	+10 to +60	118
V	5.0-6.0	570-660	710-855	4.0-4.9	2050-2080	+25 to +85	9
VI	6.0-7.5	660-790	855-1130	4.9-6.1	2060-2090	+90 to +140	5
						Total	177

The climate objective can be expressed in terms of radiative forcing (second column of Table 0.1), increase of mean temperature with respect to pre-industrial level (fifth column), concentration of greenhouse gas or carbon in the atmosphere (third or fourth column), with a certain equivalence between these measures. The decision of the climate objective then comes down to choosing between categories I to VI in Table 0.1, ordered from the most ambitious to the less ambitious objective. Despite the absence of scientific certainty about the costs of stabilization and residual damages for different stabilization levels, a political consensus tends to promote a limitation of climate warming to +2°C with respect to pre-industrial levels. This objective can be satisfied with a probability higher than 50% only with a stabilization of the concentration of all greenhouse gases at 450 ppm-CO2éq, corresponding to the more binding constraint in category I. Since concentration is now around 430 ppmCO_2-eq (against 280 ppmCO_2-eq before Industrial Revolution) and current emission trends correspond to a 2 to 3 ppmCO_2-eq yearly increase of this concentration, this objective supposes a drastic change of current trends with a fast decrease of emissions (before 2015) and a continuous pursuing of decarbonization efforts towards a 50% to 85% decrease of emissions in 2050 with respect to current levels. This is possible only of restrictive conditions are satisfied: full and immediate participation of all countries, high degree of flexibility in technical adjustments and the possibility to generate a large amount of negative emissions before 2100[6] (Krey and Riahi, 2009; van Vuuren et al, 2010). Satisfying simultaneously those conditions imposes to adopt an optimistic vision of the political, technical and behavioral barriers that may affect the patterns along which the second-best economy considered in this thesis develops.

[6]The technology envisaged to realize negative emissions is biomass-fueled power plants with sequestration, which raises questions linked to land-use competition and large-scale availability of CCS.

We adopt more conservative assumptions, which leads us to retain less extreme (but still ambitious) stabilization scenarios, corresponding to categories II and III. This means prescribing a trajectory of carbon emissions that satisfies the twofold condition: a decrease of emissions before 2030 (sixth column of Table 0.1) and around 30% decrease of carbon emissions in 2050 with respect to current levels (seventh column of Table 0.1).

3-The resource constraint

We consider here the specific role of constraints imposed by the use of scarce natural resources as production factors. In a first-best vision, explicitly accounting for these additional constraints could not but lead to an increase of mitigation costs and, conversely, a climate policy could not but increase the negative economic consequences of resource scarcity. But, if the analysis of economic interactions is extended to represent pre-existing suboptimalities, the additional constraint on resource availability can have ambiguous effects in its interactions with the climate constraint. We will focus more specifically on the effects associated to oil and urban land, which are simultaneously essential to the production process and play a crucial role in the context of climate policy because of their complex interactions with a limitation of carbon emissions.

On the one hand, climate change, energy security and the depletion of oil resources are closely related issues because of their common focus on the decline of consumption and production of this fossil energy, which is both an important source of carbon emissions, a central determinant of international trade flows and a crucial component of the energy mix. The climate policy indirectly delays the exploitation of oil, slows down its depletion and gives an early signal of the long-term scarcity of this exhaustible resource. In addition, climate policies affect the geopolitical dimension of oil markets by calling for a strategic response of major oil producers to this threat on their exportation revenues. This interplay plays a crucial role in climate negotiations as demonstrated by the claim by OPEC countries for compensations for those losses of revenues in exchange for their compliance to climate agreements (see the Article 4.8 of the UNFCCC and the article 3.14 of the Kyoto Protocol). Conversely, constraints on oil availability can be seen as positive for the long-term objective of a climate policy by limiting the amount of oil-related carbon emissions. But, these constraints may also enhance the costs of the transition towards a low-carbon economy by affecting the distribution of mitigation efforts among fossil fuels (oil, gas, coal) towards an important decrease of oil even though this source of energy is the most difficult to abstract from in the short- and medium-term due to its captive uses (transport).

On the other hand, the allocation of sites between different uses (residential, agricultural, industrial…) and the resulting land-use patterns are crucial determinants of greenhouse gas emissions in the most important sectors. We limit our analysis to the passenger transportation sector, in which location choices and infrastructure networks in the urban environment drive constrained mobility (commuting, shopping), and the freight transportation sector, in which the location of production units and consumers determines the logistics organization and the transport intensity of production/distribution processes. The mitigation potentials offered by changes in land-use patterns in these activities can be active only in the long-term because they involve changes of long-lived infrastructures. Conversely, the adoption of a climate policy favors the redirection of investments towards less carbon-intensive spatial organizations, but also limits the availability of investments for infrastructure projects at the core of these relocations because of the crowding-out effect of climate policy-related investments in the energy sector.

4. The triptych carbon price - oil price - land price

The economic consequences of a limitation of carbon emissions depend on its complex interplay with the set of other constraints that pre-exist in the economy, whether they are related to the second-best nature of economic interactions or to the limitations imposed by the use of scarce natural resources. This thesis aims at complementing existing studies of second-best economies and policies to provide an approach that encompasses all the relevant dimensions of climate policy analysis in a consistent framework. This means in particular including the specific mechanisms associated with oil and land as non-renewable resources at the heart of climate issues.

This leads us to reconsider the analysis of the costs of a climate policy, co,ventionnaly focused on the effects of the carbon price as the essential driver of the increase of end-use energy prices affecting households' purchase power and firms' production costs. However, this vision neglects the feedback effect of the climate policy on oil prices, which is yet a crucial determinant of the end-use costs of fossil energies, and the role of land prices, which measure relocation possibilities permitting a reduction of transport dependency and related emissions. To summarize, a high carbon price in a context of low oil and land prices may have less negative economic effects than a moderate carbon price applied in a very constrained context characterized by high oil and land prices.

The objective of this thesis is to propose a framework for a quantified analysis of these interactions between carbon price, oil price and land prices. This supposes giving a detailed representation of the formation of prices for non-renewable resources, which are submitted to mechanisms that are essentially different from traditional goods, in particular because of the importance of rent formation.

II- The price of scarce resources: a detour *via* the theory of rents

This section proposes an overview of the theoretical debates about the notion of rent, which will serve to identify the major mechanisms at play in the formation of scarce resource prices.

1. The birth of the notion of rent

The concept of rent appears in economic thought with the concept of the "net product", at the heart of physiocracy. This theory, developed by French economists during the 18st century, was funded on the idea that all wealth comes from agricultural production, consistently with the dominant role of this sector in the economy at that time. In this vision, land is the source of value creation and the "net product" represents total agricultural production minus the cost of past investments and the payment to farmers cultivating land. This notion plays a crucial role in the "Economic Table", invented by Quesnay in 1759, which constitutes the first attempt to represent consistent economic flows between different types of economic agents. In this primitive version of input-output tables formalized by Leontief in 1930, the economy is reduced to a unique production sector – agriculture – and three types of agents: the "Productive" class consisted of all agricultural laborers, the "Sterile" class made up of artisans and merchants and the "Proprietary" class consisted of landowners. Landowners perceive the "net product" from farmers in exchange for the right to cultivate the land they own in accordance with their "natural rights". This source of income, which appears as crucial to ensure the closure of economic flows, is not directly related to production (landowners do not participate to the production process if not by allowing farmers to work on their land) and then constitutes remuneration that has certain similarities with what will be later called rent.

The first explicit reference to the notion of rent appears in the analysis of income distribution by Adam Smith. Rent is there considered as the third source of income to be considered in parallel with wages perceived by workers and profits of capital owners. In line with previous analyses of the agricultural sector, land rent is formulated as a driver of prices, the possession of land giving right to remuneration similarly to the possession of capital. The existence of rents then depends on the use of the land, only food production ensuring systematically a revenue to landowners:

> « Human food seems to be the only produce of land which always and necessarily affords some rent to the landlord. Other sorts of produce sometimes may and sometimes may not, according to different circumstances. » (Smith, 1776, chapter 11)

This specific role of food production in rent formation must be put back in the context of an essentially agricultural economy, in which land is the only factor of production ensuring the production of more wealth that the total of what has been consumed in the production process. Economic theory will long attribute a specific role to agricultural production in rent formation, but the appearance of new dominant types of production will lead to extensions of this notion and the appearance of rents will be more generally associated to the use of a production factor that is submitted to a certain form of scarcity. Mining resources, at the core of energy-intensive production patterns of the Industrial Revolution, are an emblematic example:

> « Whether a coal-mine, for example, can afford any rent depends partly upon its fertility, and partly upon its situation. A mine of any kind may be said to be either fertile or barren, according as the quantity of mineral which can be brought from it by a certain quantity of labour is greater or less than what can be brought by an equal quantity from the greater part of other mines of the same kind. » (Smith, 1776, chapter 11)

The existence of a rent is then conditioned by the characteristics of the resource under exploitation when compared to similar resources, which constitutes the basis of the differential theory that will be formalized later by Ricardo.

2. Rent: cause or consequence of prices?

In a first approach, Adam Smith considers that the amount of rent reflects the maximum level that landowner can obtain from the farmers. It corresponds to a tradeoff on the revenue perceived by the landowner, which must be sufficiently high to provide an incentive to put their land in exploitation, but sufficiently low to ensure that farmers have enough revenues to live once this transfer to landowner has been done. In this case, the level of rent is associated to a situation of monopoly and will be a determinant of the final price of the good:

> «The rent of land, therefore, considered as the price paid for the use of the land, is naturally a monopoly price. It is not at all proportioned to what the landlord may have laid out upon the improvement of the land, or to what he can afford to take; but to what the farmer can afford to give. » (Smith, 1776, chapter 11)

In a second approach, rent is a source of income essentially different from wages and profits because it can be modified without affecting directly production. This means that rent is a consequence of prices, which appears only when considering resources creating a residual revenue (or surplus) above production costs so that high rent is conditioned by high prices

> « Rent, it is to be observed, therefore, enters into the composition of the price of commodities in a different way from wages and profit. High or low wages and profit are the causes of high or low price; high or low rent is the effect of it. » (Smith, 1776, chap 11)

Malthus elaborates on this second interpretation of rent as a surplus over production costs by pointing the necessity to analyze the reasons for high prices of commodities. These must be distinguished from monopoly prices associated to supply restrictions since the appearance of high commodity prices is driven by the constrained demand for those essential goods:

> « by applying occasionally the term monopoly to the rent of land, without stopping to mark its more radical peculiarities, [Adam Smith] leaves the reader without a definite impression of the real difference between the cause of the high price of the necessaries of life, and of monopolized commodities. » (Malthus, 1815)

Malthus then puts forward a three-step mechanism explaining high prices and hence rent formation for those essential goods. First, the quality of the soil as a production factor is a prerequisite for the appearance of rents since it allows for a production level that exceeds the amount necessary to pay farmers. Second, a demand exists for this surplus over production costs because of the basic needs for these essential goods. Third, fertile land is scarce and the increase of demand with demographic growth imposes to exploit less and less fertile land. This rise in demand drives an increase of the exchange value of the good and hence of its price, which enhance the surplus obtained on the more fertile land still in operation. This analysis conducts Malthus to identify four sources of rent formation, which combine the decrease of production costs and the increase of the exchange value of the good: capital accumulation decreasing their return, population increase decreasing wages, technical progress improving the efficiency of the production process and demand intensification fostering an increase of the final price.

This fragmented vision of rents – sometimes cause, sometimes consequence of prices – was a symptom of the necessity to develop a theoretical framework for thinking this notion. This is the differential surplus theory, whose fundamentals were already present in the above analysis, but which was consistently formalized by David Ricardo.

3. The differential theory of rents and extensions

The fundamentals of rent in David Ricardo's theory are close to the approach developed by previous authors. It is indeed justified by the inherent characteristics of the land:

> « [rent is] that *portion* of the product of the earth which is paid to the landlord for the use of original and *indestructible* power of the soil » (Ricardo, 1817, chapter 2)

However, Ricardo can be distinguished by his successful attempt to go beyond the analyses that attribute the "power of the soil" to a gift from nature by describing it instead as an economic phenomenon resulting from the interplay between land and labor. Indeed, the quality of the land is not limited to its intrinsic characteristics but is measured by its productivity, namely the quantity of labor necessary for the production process on this land. This means that agricultural products are produced from a human effort measured by the quantity of labor incorporated in the good. The more fertile the land, the higher the production level for the same quantity of capital and labor (or, equivalently, the lower the labor needs for the same production level), and hence the lower the production costs. Price formation obeys to a law that prefigures marginalist principles: price is set at the production cost on the less fertile land under exploitation. As a consequence, the possessors of the more fertile land benefit from a surplus over their production costs, which is identified as a rent: rent is a consequence of prices. This approach, which combines the theory of incorporated value, marginalist principles and rents is known as the "differential theory of rents". It allows overcoming some apparent contradictions of previous approaches, like for example explaining why, contrary to Smith's assertion, land may offer no rent in case of low population density. Indeed, in this case, only the more fertile lands are exploited because they are sufficient to satisfy the whole demand so that prices remain low, at the level corresponding to production costs over these very productive lands. On the contrary, rent appears under the pressure of increasing demand, which forces extending production to less fertile lands:

> « It is only, then, because land is not unlimited in quantity and uniform in quality, and because in the progress of population, land of an inferior quality, or less advantageously situated, is called into cultivation, that rent is ever paid for the use of it. » (Ricardo, 1817, chapter 2)

The differential theory also permits to unify approaches previously developed by Adam Smith, who arbitrarily distinguished rents resulting from land use and those arising from mines exploitation, the former being supposed to be the only type ensuring systematically a rent. By generalizing the differential principles to any natural resource with heterogeneous characteristics, Ricardo demonstrates that mining resources can lead to the emergence of rents in a similar manner than agricultural land. The only reason why it may not be the case is a low demand for mining products in the agricultural-dominated economy that Adam Smith experienced, which makes the exploitation of less fertile mines not necessary.

This unification of the rent theory for land and all other resources is formulated by John Stuart Mill, who introduces the idea that land is a productive capital, admittedly particular, but which must follow the standard laws of returns to production factors. To this aim, he distinguishes two types of land revenues: the land rents associated to its intrinsic characteristics like in Ricardo's approach *vs.* the revenues permitted by additional capital investments on this land. In some cases, this latter category must be considered as a rent, especially when they result from important investments aimed at improving long-term land productivity:

> « with regard to capital actually sunk in improvements, and not requiring periodical renewal, but spent once and for all in giving the land a permanent increase of productiveness, it appears to me that the return made to such capital loses altogether the character of profits, and is governed by the principles of rent.» (Mill, 1871, p. 408.)

This means that economic agents can influence the amount of rents, and in particular that rents can emerge from uniformly fertile lands, provided heterogeneous distribution of installed capital. Here, land is not the source of rents by itself, which is rather due to its heterogeneous characteristics as a production factor. This approach opens the way to a generalization of the differential theory to all production factors without limitations to land:

> « All advantages, in fact, which one competitor has over another, whether natural or acquired, whether personal or the result of social arrangements, assimilate the possessor of the advantage to a receiver of rent. » (Mill, 1871, p 459)

4. Scarcity rents

The differential theory of rents considers that the value of a good depends on the quantity of labor incorporated in its production, including the amount corresponding to capital investments necessary for production, and is therefore focused on the supply-side characteristics as determinants of prices and hence of rents. The marginalist approach, initiated by Jevons, Menger and Walras, adopts a complementary vision of price formation, focused on the demand-side of economic interactions. According to this theory, the exchange value of a good depends on the utility it provides, which, according to the decreasing marginal utility principle, is even higher than consumption levels are low. This means in particular that scarce goods are characterized by high prices, which are not due to their production costs but rather to the limits on their availability enhancing competition between agents. The possessor of these scarce goods then benefit from rents in the form of surplus over their production costs. This suggests a reinterpretation of the notion of rents at the basis of the notion of "scarcity rent", which pays the services offered by the good as measured by the utility it provides. These two approaches – differential rents and scarcity rents – may seem contradictory but have been reconciled by Marshall, for whom it is only two aspects of the same problem know as "economic rent":

« In a sense all rents are scarcity rents, and all rents are differential rents. But in some cases it is convenient to estimate the rent of a particular agent by comparing its yield to that of an inferior (perhaps a marginal) agent, when similarly worked with appropriate appliances. And in other cases it is best to go straight to the fundamental relations of demand to the scarcity or abundance of the means for the production of those commodities for making which the agent is serviceable. » (Marshall, 1890, BookV, Chapter 9)

The major innovation introduced to bridge the gap between these two notions is to consider the temporal dimension as a determinant of the formation of rents. The selling prices of goods and hence the surplus over production costs indeed depends on the temporal horizon considered:

« The conditions which govern the amount of this surplus and its relations to value depend not so much on the nature of the industry as on the period of time for which the calculation is made. » (Marshall, 1893)

In the short term, there exists a period of time over which the producer cannot modify his production capacities because of inertias of production systems. These supply-side constraints drive a temporary increase of prices and create temporary scarcity rents or "quasi rents", which disappear after the long term adjustments of the production processes. Over the long term, only remain the differential rent arising from the surplus over production costs. Here, in addition to the above identified determinants of rents (intrinsic characteristics of land and productivity improvements due to capital investments), Marshall adds a third component encompassing all elements that can justify a difference between prices and the value incorporated in the produced good. It can be either the relative scarcity created by an intensification of demand (for example due to population growth) or the external advantages associated to the environment or the localization (proximity to markets, intra-industry agglomeration effects, accessibility, transport costs). These latter dimensions affect the value of land because they define its production potentials over a multiplicity of potential uses, and restricting the use of a given plot of land to a specific type of production imposes to pay it at a sufficiently high level to compensate for the advantages that it could provide if devoted to other uses:

« The full rent […] is made up of three elements; the first being due to the value of the soil as it was made by nature; the second to improvements made in it by man; and the third, which is often the most important of all, to the growth of a dense and rich population and to facilities of communication by public roads, railroads, etc. » (Marshall, 1890, Book IV, Chapter 3)

Synthesis: Rents and price formation

We will understand rents as the payment of a production factor in excess with respect to the minimum level to obtain the expected service from it. As we have seen, different approaches have been proposed to explain the mechanisms driving the emergence of these rents, their amount and their distribution. We will retain the twofold distinction:

- differential rent *vs.* scarcity rent. The former focuses on the supply-side specificities as sources of rent formation in the cases where production costs are heterogeneous, whereas the latter insists on the specificities of demand in terms of competition for the possession of goods providing a high level of utility. This opposition refers to different visions of economic adjustments, either dominated by supply-side or demand-side dimensions.
- surplus rent vs. monopoly rent. The former approach attributes the emergence of rents to the excess of prices over production costs and defines them as a consequence of prices decided by market interactions. The latter considers instead that producers have the possibility to act strategically to force an increase of prices and the will to capture rents is then a cause of the price increase.

We do not consider these distinctions as revealing internal contradictions, but consider instead that they offer complementary answers to the two fundamental questions associated to markets where rents play an important role: what mechanism drives the formation of price? What is the equilibrium price level?

III- Outline of the thesis

The thesis is organized around six chapters. This succession follows the parallel elaboration of a modeling framework embarking the constraints identified in section I and II, including those imposed by oil and urban land, and of new policy-relevant insights on the interplay between climate policy, oil markets and urban land use. Note that the chapters can also be read independently, each of them being conceived to provide an autonomous vision of a distinctive step of the analysis.

Part A considers the fundamentals of the interactions between long-term oil markets, socio-economic trajectories and climate policies under technical and resource constraints. ***Chapter 1*** describes the modeling assumptions adopted to represent inertias on oil markets in a modeling framework for the assessment of climate policy under second-best economy. This framework is used to analyze oil markets and their macroeconomic consequences on both oil-exporting and oil-importing economies at different time horizons. This analysis serves to discuss the rationale of different geopolitical decisions from major oil exporters and hence to delineate the future of the interplay between oil markets and the macroeconomy. ***Chapter 2*** introduces the climate policy, and discusses its adverse impact on major oil producers, in the form of exportation revenue losses and slowing down of domestic macroeconomic growth. This analysis serves as a basis for the discussion of geopolitical dimensions of the interplay between oil markets and a climate policy, with a focus on the issue of monetary compensations and the actual participation of major oil producers to the climate coalition. ***Chapter 3*** considers more generally the time profile of mitigation costs consecutive to the implementation of an ambitious climate policy at the world level over the period 2010-2100. It identifies in particular the role of oil markets in driving these profiles in the form of both high risks of important short-term losses and opportunities of co-benefits ensuring relative gains thanks to the correction of imperfections in baseline scenarios. This chapter concludes by demonstrating the role of complementary policies to carbon pricing that help decreasing mitigation costs by reducing the dependence on mobility and even create room for negative mitigation costs over the long term.

Part B introduces location patterns and the spatial dimension of economic activities as a crucial determinant of energy and environmental effects. ***Chapter 4*** proposes a theoretical model of the interplay between economic activity, location decisions and climate change in the light of the New Economic Geography. This is done by representing the stock effect of pollutants affecting welfare in function of their accumulation over time and the potentials for catastrophic event at low environmental quality in the long-term. This model serves as a basis for an analysis of the interplay between trade patterns, regional spatial structure and long-term stabilization objectives of pollutant stocks. ***Chapter 5*** develops a model of the interactions between a multiplicity of agglomerations, their interplay being described according to NEG principles and their internal structure (including the price of urban land) being represented in line with urban economy theory. This framework is calibrated on explicit datasets for the 74 largest OECD agglomerations, its dynamics is tested against past historical data and it represents the specificities of agglomerations driving urban dynamics. ***Chapter 6*** embarks this analysis of location choices into the energy-economy model developed in Part A for the analysis of climate change issues. This allows representing both the impact of macroeconomic trends on urbanization patterns and, conversely, the feedback effect of urban development on the macroeconomy. This framework is used to reassess the costs of climate policies and to evaluate in particular the potentials offered by densification policies at the urban scale as a mean to reduce mobility dependency and related carbon emissions.

References

Clarke L, Edmonds J, Krey V, Richels R, Rose S and Tavoni M (2010). "International Climate Policy Architectures: Overview of the EMF 22 International Scenarios". Energy Economics 31 (Supp. 2): 64-81.

Combet E, Ghersi F, Hourcade JC and Thubin C (2010). Economie d'une fiscalité carbone en France. CFDT/ADEME/ IRES Report, 141p.

Crassous R (2008). Modéliser le long-terme dans un monde de second rang : applicationaux politiques climatiques. PhD thesis.

Edenhofer O, Lessmann K, Kemfert C, Grubb M and Köhler J (2006). "Induced technological change: Exploring its implications for the economics of atmospheric stabilization. Synthesis report from the innovation modeling comparison project". *In*: Endogenous Technological Change and the Economics of Atmospheric Stabilisation. The Energy Journal (Special Issue 1): 57-108.

Edenhofer O, Knopf B, Barker T, Baumstark L, Bellevrat E, Chateau B, Criqui P, Isaac M, Kitous A, Kypreos S, Leimbach M, Lessmann K, Magné B, Scrieciu S, Turton H and van Vuuren DP (2010). "The economics of low stabilization: model comparison of mitigation strategies and costs". *In*: [Edenhofer O, Knopf B, Leimbach M, Bauer N (Eds)] The economics of low stabilization. Energy Journal 31(Special Issue 1): 11-48

Goulder L (1995). "Environmental Taxation and the "Double Dividend": A Reader's Guide". NBER Working Paper 4896, 27 p.

Guivarch C(2010). Evaluer le coût des politiques climatiques : de l'importance des mécanismes de second rang. PhD thesis.

Hourcade JC, Jaccard M, Bataille C, Ghersi F (2006). Hybrid Modeling: New Answers to Old Challenges. The Energy Journal, Special Issue 2006, 1-12..

IPCC (2007). Summary for Policymakers. *In*: [Metz B., Davidson OR., Bosch PR, Dave R, Meyer LA (Eds)] Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jakob M, Luderer G, Steckel J, Bosetti V, Tavoni M and Monjon S (2011). "Time to act now? Assessing the costs of delaying climate measures and benefits of early action". *In*: The Economics of Ambitious GHG Emission Reductions**.** Climatic Change (*forthcoming*).

Krey V and Riahi K (2009). “Implications of delayed participation and technology failure for the feasibility, costs, and likelihood of staying below temperature targets—Greenhouse gas mitigation scenarios for the 21st century”. *In*: International, U.S. and E.U. Climate Change Control Scenarios: Results from EMF 22. Energy Economics 31 (Supp.2): 94-S106

Ligthart JE (1998). “The Macroeconomic Effects of Environmental Taxes: A Closer Look at the Feasibility of “Win-Win“ Outcomes”. IMF Working Paper 98/75.

Luderer G, DeCian E, Hourcade JC, Leimbach M, Waisman H and Edenhofer O (2011). On the regional distribution of mitigation costs in a global cap-and-trade regime. *In*: The Economics of Ambitious GHG Emission Reductions. Climatic Change (*forthcoming*).

Malthus (1815). An Inquiry into the Nature and Progress of Rent, and the Principles by which it is Regulated.

Marshall A (1893). “On Rent”. Economic Journal 3: 74-90.

Marshall A (1890). Principles of Economics (7th edition). London: Macmillan & Co.

Mill JS (1871). Principles of Political Economy (7th edition). London: Cooperative Publishing Society.

Ricardo D (1817). On the Principles of Political Economy and Taxation (3rd edition). London: John Murray,

Sassi O (2008). L'impact du changement technique endogène sur les politiques climatiques. PhD thesis.

Smith A (1776). An Inquiry into the Nature and Causes of the Wealth of Nations (5th edition). London: Methuen and Co.

Stern N (2007). The Economics of Climate Change: The Stern Review. Cambridge University Press, Cambridge, 692 pp.

Tavoni M, De Cian E, Luderer G, Steckel J and Waisman H (2011). “The value of technology and of its evolution towards a low carbon economy”. *In*: The Economics of Ambitious GHG Emission Reductions**.** Climatic Change (*forthcoming*).

van Vuuren DP, Stehfest E, den Elzen MGJ, van Vliet J and Isaac M. (2010) “Exploring IMAGE model scenarios that keep greenhouse gas radiative forcing below 3 W/m^2 in 2100 ”. Energy Economics 32 (5):1105-1120

Weyant JP (ed.) (2004). EMF 19 alternative technology strategies for climate change policy. Energy Economics, 26(4): 501-755.

Weyant JP, de la Chesnaye F and Blanford G (2006). "Overview of EMF– 21: Multigas mitigation and climate policy". *In:* Multi-Greenhouse Gas Mitigation and Climate Policy. The Energy Journal (Special Issue No.3): 1-32.

Yohe GW, Lasco RD Ahmad QK, Arnell NW, Cohen SJ, Hope C, Janetos AC and Perez RT (2007) Perspectives on climate change and sustainability. *In:* [Parry, ML, Canziani, OF, Palutikof, JP, van der Linden, PJ and Hanson, CE (Eds)] Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK: 811-841.

Part A

Oil markets and climate policies

This introductory section derives lessons from the general analysis of rents (see General Introduction) to define the modes of representation for oil markets that will be adopted in this thesis

I- Short- and long-term scarcity rents

The formation of oil prices is governed by the limitations on the deployment of production capacities and on the potentials for reducing demand, the magnitude of these two constraints depending on the temporal horizon considered.

In the short-term, oil markets combine a very inelastic demand with important constraints on the deployment of production capacities that make oil rents belong to the category of "quasi rents" in the sense of Marshall (1890). On the one hand, the short-term inelasticity of oil demand is notably due to the dominant role of the transport sector (52% of world oil demand in 2006 according to (IEA, 2008)), which imposes specific constraints in reason of (a) the low potentials for substitution towards other energy sources that limit the decoupling between oil demand and transport activity, since other potential sources of liquid fuels (biofuels and coal liquefaction) are not immediately available at a large scale, and (b) the dependence on mobility that cannot be reduced overnight because of the constraints imposed by long-lived infrastructure. On the other hand, the deployment of oil production capacities is submitted to strong inertias associated to the high costs of findings and developments (around 20$/Barrel according to Figure 13.6 in (IEA, 2008)) and the delays before production can effectively start. This means that the potentials for production increase are economically and technically limited in the short-term.

In the long run, the existence of oil rents is more fundamentally due to its exhaustible nature since oil is a capital given by nature that humans cannot produce. Moving from a resource which can be consumed without affecting its capital (land) to an exhaustible resource whose consumption at the current date affects the availability in the future (oil) makes necessary the representation of intertemporal effects. This dimension has been treated by Hotelling (1931), who proposes a theoretical framework for analyzing the behavior of producers exploiting an exhaustible resource. He demonstrates that the production profile that optimizes producers' profits is obtained by an exponential increase of prices and a progressive decline of supply. The resulting rent is essentially different from a differential rent since it exists even with a homogenous resource with zero production costs.

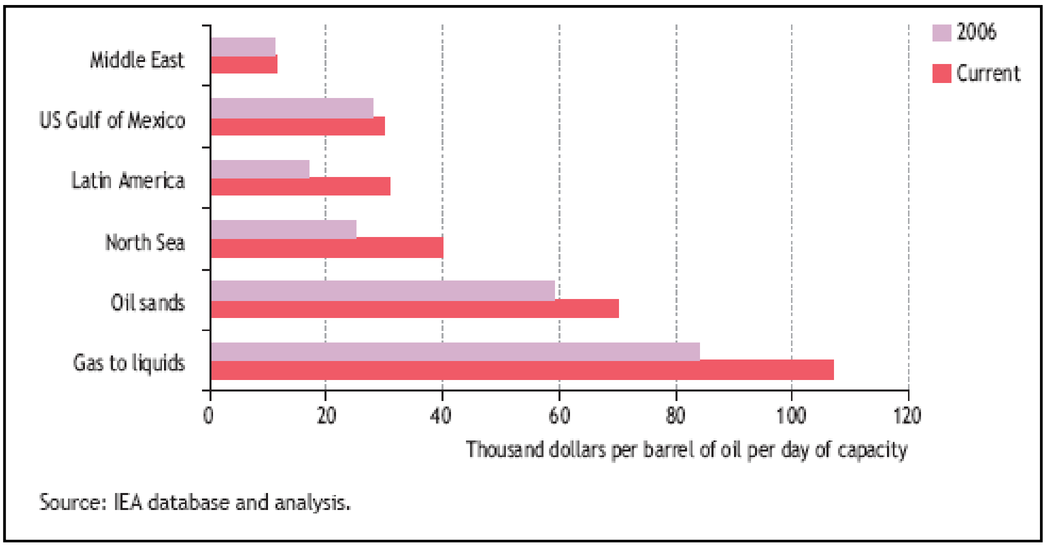
II- Resource heterogeneity and differential rents

When accounting for the heterogeneity of a natural resource, differential rents can emerge as illustrated by the direct parallel with the mechanisms at play with lands of different fertilities established by Ricardo on the example of mines:

> «There are mines of various qualities, affording very different results, with equal quantities of labour. […] The return for capital from the poorest mine paying no rent, would regulate the rent of all the other more productive mines. […] Since this principle is precisely the same as that which we have already laid down respecting land, it will not be necessary further to enlarge on it. » (Ricardo, 1817, chapter 3).

The example of mines can be directly transferred to oil resources, since intrinsic differences of oil fields (deepness, accessibility, location, nature of the resource) introduce differences of production costs, which play the role of the "quality" of the resource. This heterogeneity is illustrated in Figure 0.1, which reports the capital costs necessary for putting different types of oil resource in exploitation. It demonstrates the high difference in the cost of putting a barrel on the market according to the regions and the types of oil under consideration. For example, upstream investments are four times higher in the North Sea than in Middle-East and even seven times higher when extending the range of oil resources to non-conventional oil like oil sands.

Figure 0.1: *Average capital cost of upstream projects under development (IEA, 2008, Figure 13.10)*



However, these differences are the source of differential rents only if demand is too strong to be satisfied only by the low-cost categories. The empirical analysis confirms that this is actually the case in the real world, since low-cost oil resources (principally located in Middle-East regions) only contribute to a minority of total production, the rest of demand being satisfied by higher-cost categories (Table 0.2)

Table 0.2: *Oil production in 2007, in MBarrel per day (source: IEA, 2008)*

Middle-East	Rest of OPEC	North America	Rest of OECD	Latin America	Rest of the World	TOTAL
23.7	12.2	13.8	5.5	3.5	25.6	84.3

This production scheme may seem surprising given the abundance of low-cost reserves that could potentially meet the whole demand, but the simultaneous exploitation of different oil categories is due to number of specificities that limit the expansion of existing capacities: (a) under given technology, the extraction rate tends to decrease with the exploitation of an oil field so that additional investments triggering an increase of production costs are necessary to maintain the flow of production. This means that the exploitation of low-cost categories is less favorable when depletion approaches and it may be appropriate to turn to other categories which, although initially more expensive, are less advanced in their depletion process, (b) the limited quantity of low-cost categories is limited and the exploitation of high-cost categories is unavoidable. When accounting for the already mentioned inertias in the deployment of new production capacities, the producers of the high-cost categories must anticipate their entrance into the market and start production, and (c) the interests of producers in different regions of the world pushes even the possessors of high-cost categories to put their reserves into exploitation to benefit from a small amount of rent without waiting for their categories to be the more profitable remaining resource

III- Market power and monopoly rents

Finally, major oil producers grouped in the OPEC have both the possibility and the will to control production in order to satisfy their rent-seeking objective. This is possible since they benefit from a high market share, which is likely to be reinforced within the next decade because of the decline of conventional production in other regions of the world. Although the coordination of this oligopoly is sometimes erratic, it remains that this group of suppliers have the possibility to act strategically to influence oil prices through their decisions on the deployment of production capacities, as typically demonstrated by the oil shocks of the 70's resulting from sudden changes of OPEC's production decisions. In this case, the oil rent has a strong "monopoly rent" component.[1]

IV. Specifications for the representation of oil markets

The superposition of the different types of rents above described is at the core of the representation of oil prices. Given the above analysis, their fundamentals can be summarized as follows:

- On the demand side, the main determinants to be considered are those explaining the low elasticity at short- and long-term as a consequence of the interplay between technical constraints (availability of alternatives to oil, limits on the decoupling between oil demand and economic activity), consumption patterns (preferences) and location choices (constrained mobility)
- On the supply side, it is necessary to represent the short-term inertias on the expansion of production capacities, the long-term limitations imposed by resource depletion, the heterogeneity of the oil resource in terms of regional distribution and production costs, and OPEC's market power, which let them affect world prices through their decisions of production expansion.

Chapter 1 describes and discusses the modeling choices adopted to capture these specificities of oil markets. This architecture will serve in the remaining of this thesis as a tool to investigate the interplay between oil markets and climate policy. Chapter 2 focuses on the adverse impacts of climate policies on major oil exporters and discusses the crucial geopolitical issue of monetary compensations. Chapter 3 investigates the cost of climate policies with in particular the co-benefit in terms of reduced vulnerability to oil scarcity.

[1] It must be noted that the 2008-2009 rise of oil prices, commonly considered as the third oil shock, cannot be attributed to such a rent-seeking behavior, but rather belongs to the scarcity rent category described in section I. It was indeed due to the very fast increase of oil demand consecutive to economic growth in emerging economies and to the impossibility to deploy immediately sufficient production capacities.

REFERENCES

Hotelling H (1931). "The Economics of Exhaustible Resources". Journal of Political Economy 39:137-175.

IEA (2008). World Energy Outlook 2008. International Energy Agency /OCDE, Paris

Marshall A (1890). Principles of Economics (7th edition). London: Macmillan & Co.

Ricardo D (1817). On the Principles of Political Economy and Taxation (3rd edition). London: John Murray.

Chapter 1

Peak Oil through the lens of a general equilibrium assessment

This first chapter[*] examines the interactions between the determinants of oil markets – oil supply, fuel demand and oil substitutes – and disentangles their impact on the time profile of oil prices and the macroeconomy. To this aim, we elaborate a general equilibrium model in which production and prices endogenously emerge from the interplay between the technical, macroeconomic and geopolitical determinants of oil markets under non-perfect expectations. This allows in particular to represent Peak Oil profiles and their macroeconomic effects at different time horizons. We consider two oil price profiles corresponding to alternative strategies of Middle-East producers, which prove to have weak influence on the date of Peak Oil but important macroeconomic effects on OECD and Middle-East growth trajectories. To test the robustness of our findings, we investigate Middle-East's trade-off for different objectives (maximisation of oil revenues or households' welfare) and a large set of assumptions about conventional oil resources and deployment of non conventional oil.

[*] This chapter is a reproduction of : Waisman H, Rozenberg J, Sassi O and Hourcade JC (2011). Peak Oil through the lens of a general equilibrium assessment, *submitted to* Energy Policy.

In public debates, Peak Oil relays concerns about the date at which world oil production will start declining inexorably. The debates have been focused on the date of this Peak Oil and are essentially conducted under the assumption that oil production profiles are determined by exogenous assumptions on the total amount of oil resources (see (Al-Husseini, 2006) for a review). This vision is supported by the generalization, at a global level, of bell-shaped profiles used by Hubbert to predict the decline of US production in the 70's ((Hubbert, 1956, 1962); Deffeys (2002)). Note that these curves are meant to capture geological constraints in the form of depletion effects and inertias on the deployment of production capacities.

This paper finds its starting point in the idea that the focus on the geological origin and the date of Peak Oil distracts the attention from the core determinants and the economic consequences of the end of cheap oil. Setting aside controversies about the generalization at a macro level of the Hubbert approach (Lynch, 2003), this paper argues that what matters is not so much the date of Peak Oil than the abruptness of the unanticipated break in oil trends at that period and the capacity of the economies to adapt to it.

This abruptness and its economic consequences are determined by the relative evolution rates of oil supply, fuel demand and oil substitutes under imperfect expectations and inertia constraints. To investigate the interplay between these dimensions, we use a Computable General Equilibrium (CGE) model, which incorporates a comprehensive description of the determinants of oil markets, including the geological constraints behind the Hubbert curves. This framework pictures a world with imperfect foresight, endogenous technical change and inertia on the deployment of end-use equipments and oil substitutes. Section 1 describes and justifies this modeling option.

Section 2 conducts a comparative analysis of the economic consequences of two oil pricing trajectories: high short-term prices caused by a limited deployment of production capacities *vs.* moderate short term prices caused by a market flooding behavior. The former allows high short-term revenues for oil-producing countries, while it limits the vulnerability of oil-importing economies to Peak Oil by accelerating oil-free technical change; the latter discourages oil-saving technical change and triggers high prices in the Peak Oil period.

Section 3 conducts a sensitivity analysis on the results by considering different assumptions regarding the amount of oil resources and the extent of inertias that characterize non-conventional production. We assess their impact on economic outcomes and show in particular the parameter sets under which the temporary sacrifice of short-term oil profits under the market flooding option may prove beneficial for Middle-East producers thanks to the later increase of their revenue.

I. Endogenizing Peak Oil in a second-best economy

Long run general equilibrium interactions between oil markets and economic growth are conventionally investigated either with models picturing exhaustible resource exploitation *à la* Hotelling (1931) which conclude, instead of a Peak Oil, to a steady decline of production over time (see, for example, Anderson (1972), Solow (1974) or Stiglitz (1974) and Krautkraemer (1998) for a review)[1], or with energy-economy models which conventionally assume steady growth pathways and aggregate supply curves (IPCC, 2007). With these approaches, meant to explore long run pathways, the geological constraints on short term adaptability of oil production do not really matter because they are anticipated and/or because the oil demand, driven by steady growth, evolves smoothly.

The short-term effects are considered through two independent traditions. On the one hand, econometric analyses developed after the oil shocks investigate the transmission channels between oil prices and GDP but do not account for long term resource depletion because of their short-term focus (Hamilton (2008)). These studies demonstrate that modeling exercises can better reproduce the observed magnitude of the economic effect of oil price variations if they include *1) mark-up pricing* to capture market imperfections (Rotemberg and Woodford, 1996); *2) partial utilization rate of capital* when the full utilization of installed production capacities cannot be achieved due to limits in the substitution between capital and energy (Finn, 2000); *3) a putty-clay description of technologies* to represent the inertias in the renewal of capital stock (Atkeson and Kehoe, 1999); *4) frictions in the reallocation of capital across heterogeneous sectors* causing differentiated levels of idle production capacities (Bresnahan and Ramey, 1993); *5) frictions in the reallocation of labor across heterogeneous sectors* causing differentiated levels of unemployment (Davis and Haltiwanger, 2001). On the other hand, recursive partial equilibrium analyses of supply/demand adjustments can predict Peak Oil but fail to consider their macroeconomic impacts (see (Fattouh, 2007) for a review). This group of studies teaches us the crucial role played by geological constraints, geopolitical dimensions, technical inertias and imperfect foresight on short-run oil supply adaptability.
The CGE model IMACLIM-R bridges the gap between these different branches of the literature by capturing the general equilibrium effects of short-term dynamics in second-best economies at different time horizons.

[1] A notable exception is in Holland (2008) who obtains a peak of production in an Hotelling-like framework by embarking forces that increase the equilibrium production and counterbalance the decreasing trend imposed by the depletion effect

1- Modeling the impact of oil markets on macroeconomic dynamics

IMACLIM-R is a recursive CGE model of the world economy, divided in 12 regions and 12 sectors (see Annex A for technical details). It is calibrated for the 2001 base year by modifying the set of balanced input-output tables provided by the GTAP-6 dataset (Dimaranan, 2006) to make them fully compatible with 2001 IEA energy balances (in Mtoe) and data on passengers' mobility (in passenger-km) from (Schafer and Victor, 2000). The model was tested against historic data up to 2006 (Guivarch et al., 2009) and covers the period 2001-2050 in yearly steps through the recursive succession of static equilibria and dynamic modules. It incorporates the above listed five features identified from econometric analyses as crucial for the representation of energy-economy interactions.

The *static equilibrium* represents short-run macroeconomic interactions at each date *t* under technology and capacity constraints. It is calculated assuming Leontief production functions with fixed intermediate consumption and labor inputs, decreasing static returns caused by higher labor costs at high utilization rate of production capacities (Corrado and Mattey, 1997) and fixed mark-up in non-energy sectors (feature 1). Households maximize their utility through a tradeoff between consumption goods, mobility services and residential energy uses considering fixed end-use equipments. Market clearing conditions can lead to a partial utilization of production capacities (feature 2) given the fixed mark-up pricing and the stickiness of labor markets (feature 5). This equilibrium provides a snapshot of the economy at date *t* in terms of relative prices, wages, employment, production levels and trade flows.

The *dynamic modules* are reduced forms of bottom-up models, which describe the evolution of structural and technical parameters between *t* and *t+1* in response to past and current economic signals. Available techniques at date *t* result from the structure and amount of cumulated learning-by-doing processes within the innovation possibility frontier characterizing explicitly the ultimate potentials on the supply and demand side (Ahmad, 1966). Technical choices modify only new input-output coefficients and not those of techniques embodied in equipments resulting from past choices. This putty-clay description helps to capture inertias on the renewal of technologies (feature 3) and capital (feature 4). Note that this description of inertia also enables a realistic reproduction of the heterogeneity in technical dynamics across regions. The new technical coefficients and investment choices are sent back to the static module in the form of updated input-output coefficients and production capacities to calculate the equilibrium at date *t+1*.

The consistency of the iteration between the static equilibrium and dynamic modules relies on 'hybrid matrices' (Hourcade et al., 2006), which ensure a description of the economy in consistent money values and physical quantities (Sands et al., 2005). This dual description represents the material and technical content of production processes and allows abandoning standard aggregate production functions, which have intrinsic limitations in case of large departures from the reference equilibrium (Frondel et al., 2002) and deep changes of production frontiers over several decades.

In this multisectoral framework with partial use of production factors, effective growth patterns depart from the natural rate (Phelps, 1961) given by exogenous assumptions on active population (derived from UN medium scenarios) and labor productivity (satisfying a convergence hypothesis (Barro and Sala-i-Martin, 1992) informed by historic trajectories (Maddison, 1995) and 'best guess' assumptions (Oliveira-Martins et al., 2005)). The structure and rate of effective growth at each point in time are endogenously determined by a) the allocation of the labor force across sectors, which is governed by the final demand addressed to these sectors b) the sectoral productivities which result from past investment decisions governing learning by doing processes c) the shortage or excess of productive capacities which result from past investment decisions under adaptive expectations.

2- Modeling the long-term dynamics of oil markets

The determinants of oil markets are described in dynamic modules which include lessons from partial equilibrium analyses of supply/demand adjustments on oil markets. They represent: the technical constraints (including geology) on the short-term adaptability of oil supply and the influence of Middle-East countries on production decisions (section 1.2.1); technical inertias on the deployment of oil substitutes (1.2.2); and consumers' short-term trade-offs in a set of technical and economic conditions (1.2.3).

2-1 Oil supply

IMACLIM-R distinguishes seven categories of conventional and five categories of non-conventional oil resources in each region. Each category *i* is characterized by the amount of ultimate resources $Q_{\infty,i}$ (given by the sum of resources extracted before 2001 and recoverable resources) and by a threshold selling price above which producers initiate production, $p^{(0)}(i)$. This price is a proxy for production costs and accessibility. Table 1.1 gives our numerical assumptions of the amount of ultimate resources in the main groups of regions. The figures are consistent with conservative estimates (USGS, 2000; Greene et al., 2006; Rogner, 1997) and a sensitivity analysis in section 3 will investigate the effect of more pessimistic or optimistic assumptions. Note that oil shales are not included because the specificities of their exploitation process and the associated high production cost lead us to consider them as an alternative to oil instead of a new category of oil.

Table 1.1. *Assumptions about oil resources in the central case (Trillion bbl)*

Resources extracted before 2001	**Recoverable resources beyond 2001***				
	Conventional oil		Non-conventional oil (Heavy oil and Tar sands)		
	Middle-East	RoW	Canada	Latin America	RoW
0.895	0.78	1.17	0.220	0.38	0.4

* « recoverable resources » are 2P reserves (Proven+Probable) remaining in the soil, which has been identified as the relevant indicator to investigate global oil peak (Bentley et al,, 2007)

Each oil category is submitted to geological constraints (inertias in the exploration process and depletion effects), which limit the pace of expansion of their production capacity. In line with (Rehrl and Friedrich, 2006), who combine analyzes of discovery processes (Uhler, 1976) and of the "mineral economy" (Reynolds, 1999), the maximum rate of increase in production capacity for an oil category *i* at date *t*, $\Delta Cap_{\max}(t,i)$, is given by:

$$\frac{\Delta Cap_{\max}(t,i)}{Cap(t,i)} = \frac{b_i \cdot \left(e^{-b_i(t-t_{0,i})} - 1\right)}{\left(1 + e^{-b_i(t-t_{0,i})}\right)} \quad (1)$$

The parameter b_i (in t^{-1}) controls the intensity of constraints on production growth: a small (high) b_i means a flat (sloping) production profile to represent slow (fast) deployment of production capacities. We retain b_i=0.061/year for conventional oil as estimated by Rehrl and Friedrich (2006) and, for the sake of simplicity, the same value for non-conventional oil in the median case (Section 3 relaxes this hypothesis by considering both lower and higher values of the *b*-parameter for non conventional oil). The parameter $t_{0,i}$ represents the date at which production capacities of the concerned oil category are expected to start to decline due to depletion effects. It is endogenous and varies in time since it depends on the amount of oil remaining in the soil given past exploitation decisions.

Non-Middle-East producers are seen as 'fatal producers' who do not act strategically on oil markets. Given the selling oil price p_{oil}, they invest in new production capacity if an oil category becomes profitable: they develop production capacities at their maximum rate of increase $\Delta Cap_{\max}(t,i)$ for least-cost categories ($p_{oil} > p^{(0)}(i)$) but stop investments in high-cost categories ($p_{oil} < p^{(0)}(i)$). If prices continuously increase, production capacities of a given oil category follow a bell-shape trend, whereas their deployment profile passes through a plateau if prices decrease below the profitability threshold.

Middle-East producers are 'swing producers' who are free to strategically time their investment decisions and, until they reach their depletion constraints, to control oil prices through the utilization rate of their production capacities (Kaufmann et al, 2004). This possibility is justified by the temporary reinforcement of their market power due to the stagnation and decline of conventional oil in the rest of the world. They can in particular decide to slow the development of production capacities below its maximum rate in order to adjust the oil price according to their rent-seeking objective.

Total production capacity at date t is given by the sum over oil categories of investment decisions which are conditioned by different production costs (captured by different $p^{(0)}(i)$ threshold). This means that projects of various merit orders coexist at a given point in time, consistently with the observed evidence and theoretical justifications[2].

2-2 Substitutes to oil

The first large-scale substitute to oil for liquid fuels production consists in first and second generation biofuels from renewable land resources. Their diffusion is controlled by supply curves borrowed from IEA (2006): at each date, biofuels' market share is an increasing function of oil prices which captures in a simplistic manner the competition between biofuels and oil-based liquid fuels (everything else being equal, the former are more competitive and

[2] (Kemp and Van Long, 1980) have indeed demonstrated that, in a general equilibrium context, the lowest-cost deposits are not necessarily exploited first. (Holland, 2003) even demonstrates that least-cost-first extraction rule does not hold in partial equilibrium under capacity constraints, like those envisaged for geological reasons here.

their penetration into the market is more prominent when higher oil price make the latter more expensive). These supply curves consider explicit limits on production due to land availability and competition with other biomass uses and are modified from one date to the other to account for learning-by-doing improvements.

The second alternative to oil is Coal-To-Liquid (CTL). We consider it as an inexhaustible backstop technology submitted to deployment capacity constraints. In line with Amigues et al (1998), production of the inexhaustible substitute starts before all the least-cost deposits of the exhaustible resource are exploited: CTL enters the market when oil prices exceed a threshold value, p_{CTL} , set for the sake of simplicity at p_{CTL}= 100$/bbl for all scenarios. Once this threshold is crossed, CTL producers are willing to fill the gap between total liquid fuel demand, *D(t)*, and total supply by other sources (refined oil and biofuels), *S(t)*. But, CTL production may be limited by constraints on delivery capacity due to past investment decisions if, due to imperfect foresight, profitability prospects for CTL were underestimated. These prospects are an increasing function of oil prices at each point in time[3] and cumulative investment on CTL over time is then a function of the sum of past oil prices: $p_{cum}(t) = \sum_{i=2010}^{t} p_{oil}(i)$. The share *s* of the potential market for CTL *D(t)- S(t)* that is actually available to CTL is thus an increasing function of $p_{cum}(t)$. As soon as oil price exceeds p_{CTL} , CTL production is then given by:

$$CTL(t) = s\left(p_{cum}(t)\right).\left[D(t) - \mathrm{S}(t)\right] \qquad (2)$$

2-3 Liquid fuels' demand

In IMACLIM-R, final demand for liquid fuels is derived from households' and industry's demand for energy services derived from utility and profit maximization respectively. Bottom-up modules describe the dynamics of technological constraints in the three major oil-consuming sectors (industry, residential, transport). Because of inertias in the renewal of end-use equipment and the pace of learning-by-doing processes, a significant decoupling between liquid fuel demand and economic growth can be obtained only after the renewal of several capital vintages, all the more so under imperfect foresight. In the transport sector, passengers' mobility and modal distribution depend on *(i)* households' choices from an explicit portfolio

[3] Indeed, higher oil prices drive higher prices of liquid fuels, including those produced from coal, and then higher profitability prospects for CTL.

of vehicles (including electric vehicles) according to minimization of the total user-costs and *(ii)* the availability and efficiency (including congestion effects) of road infrastructures and alternative options (railways, soft modes) driving the saturation of the time budget the consumer can allocate to transportation. In the long-run, the decoupling between liquid fuel demand and economic growth is constrained by *(i)* higher energy service demand (mobility, residential uses) along with wealth increase *(ii)* technical asymptotes for fuel switching and energy efficiency, *(iii)* limited potentials for non-fossil energies including political obstacles for nuclear *(iv)* increasing trends in freight mobility imposed by international trade and just-in-time processes *(v)* rebound effects in passenger mobility (Greening et al, 2000).

II. Peak Oil profiles and their macroeconomic dimensions

In this section, we study the implications of two oil pricing trajectories resulting from alternative strategic options for Middle-East producers under the same assumptions on the determinants of liquid fuel demand. We define two counterfactual scenarios[4]:

- *The Market Flooding scenario (MF)*: Middle-East producers expand their production capacities and bring the oil price back to its pre-2004 level, p_{low} = 50\$/bbl. This floor level is assumed to be sufficient to maintain the stability in the cartel and guarantees a minimum level of income to highly populated countries.
- *The Limited Deployment scenario (LD)*: Middle-East producers refrain from investing in new capacity and maintain the medium term oil price around p_{high}=80\$/bbl. They adopt local fiscal policies to secure domestic social stability by moderating the increase of energy prices for the consumers of the region. [5]

1- Beyond Peak Oil, contrasting dynamics of oil markets

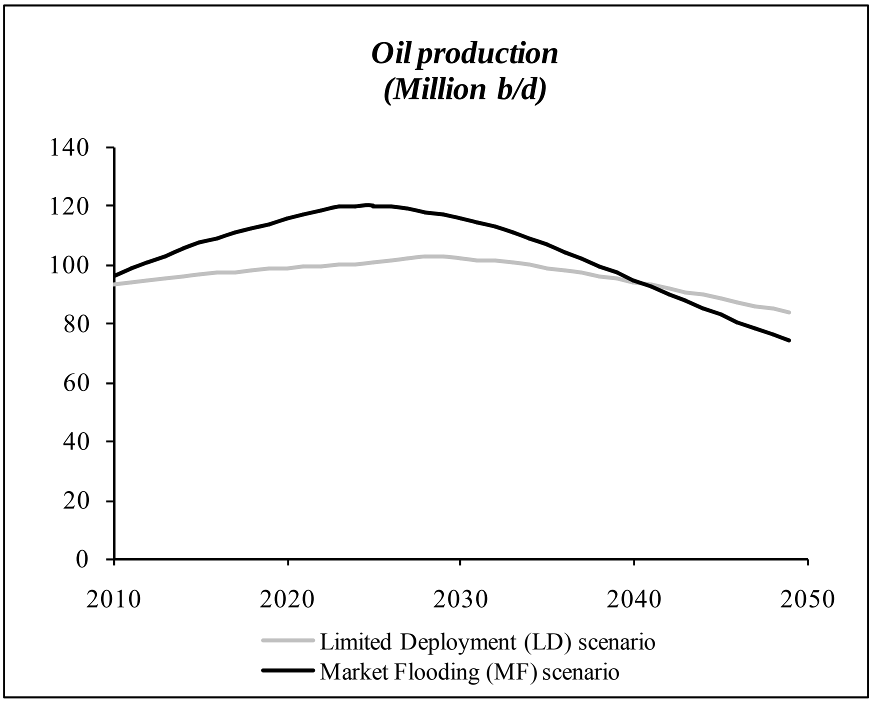
The world oil production profile proves to be bell shaped in both scenarios, peaking in 2025 in the Market Flooding scenario and in 2028 in the Low Deployment scenario (Figure 1.1). In the Market Flooding scenario, oil-intensive growth patterns are fostered by low prices which

[4] These scenarios are built on a single set of assumptions about natural growth rates, which intentionally do not represent the current economic crisis for the sake of simplicity. But, the analysis carried out in this paper provides important insights on the medium term dynamics of the economic recovery phase, which will be critically determined by the economic interactions on oil markets. Further investigation will be necessary to consider the feedback effect of the current economic crisis on the real behaviors of oil markets, specifically because of the inertia of re-launching investments in both conventional and non conventional oil.

[5] The values of p_{low} and p_{high} are expressed in 2001\$ and correspond respectively to around 60\$/bbl and 100 \$/bbl in current currency. They represent a low and high value for medium-term oil price, around the estimate of 78\$/bbl by the Short-Term Energy Outlook 2010 (available at: http://www.eia.doe.gov/emeu/steo/pub/contents.html).

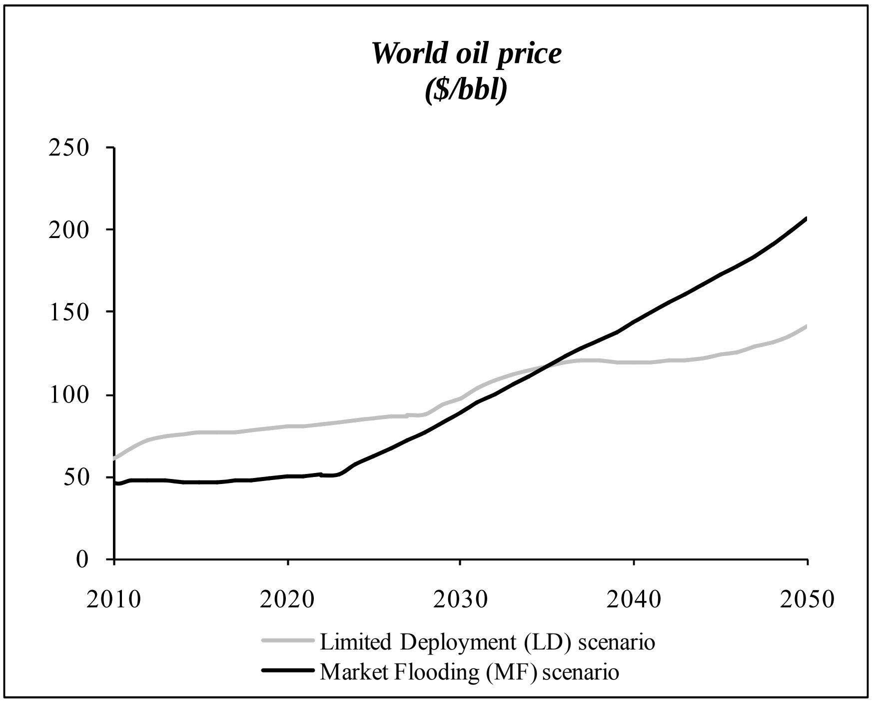
accelerate the exhaustion of conventional resources and leads to an early Peak Oil. This corresponds to a pronounced bell-shaped profile with significant break in production trends at the Peak Oil date and fast decrease after that. In the Low Deployment scenario, on the contrary, higher short-term prices foster moderation of demand and lead to a flatter profile; the reversal of production trends around the Peak Oil period is smooth and production volumes decrease at a moderate pace in the long term. Total supply even becomes higher than in the Market Flooding case after 2040.

Figure 1.1.*World oil production (Million b/d)*



The small gap in the Peak Oil dates masks indeed important differences in the production profile. The peak level is 20% higher in volume in the Market Flooding scenario (120 Million b/d) and the reversal of production trends after the Peak Oil is more abrupt (the production declines by 31% in the Market Flooding scenario and only 17% in the Low Deployment scenario over the twenty years following Peak Oil). Logically indeed, lower energy prices in the first period (a) induce intensive consumption causing faster exhaustion and sharper decline of conventional oil, and (b) deter investment in non-conventional production capacities and limit their availability in the post-Peak Oil period.

Figure 1.2. *World oil price ($/bbl)*

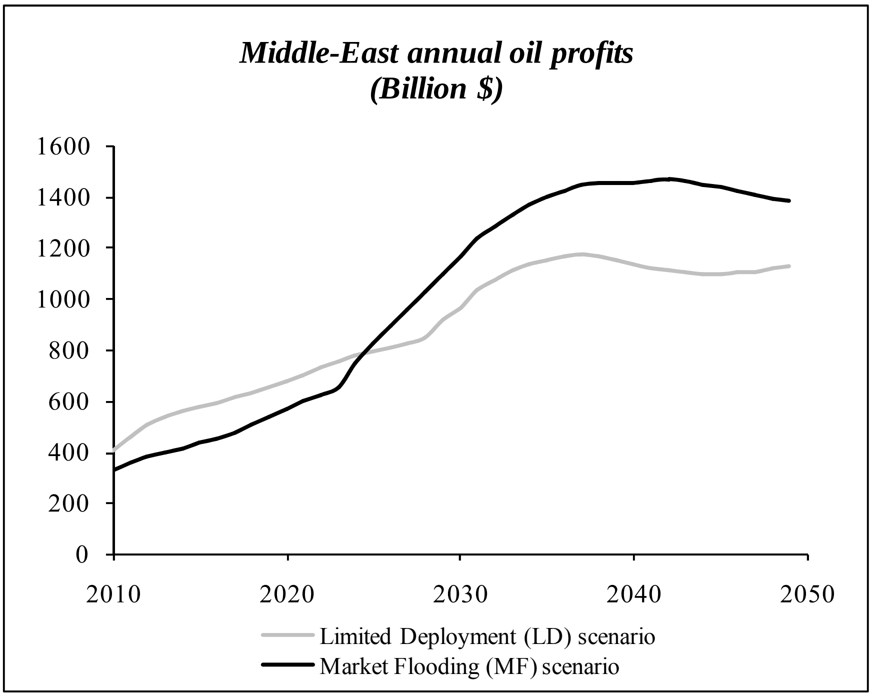


In the Market Flooding scenario, a steep and lasting surge in oil prices begins just before Peak Oil (Figure 1.2). It is triggered by tension between high demand, which cannot be reduced overnight due to inertias, and the constraints on the deployment of oil and oil substitutes' production capacities. Conversely, prices in the Low Deployment scenario increase smoothly and are lower than prices in the Market Flooding scenario after 2035, because high early price signals foster a timely penetration of oil substitutes and trigger energy efficiency abroad (Figure 2). Over the very long run, oil prices return to the price of the backstop CTL (100$/b) in both scenarios, but inertias in the penetration of this technology prevent this convergence during the period 2010-2050 considered in this paper.

2- The terms of the economic trade-off for oil producers

The time-profile of Middle-East oil profits (Figure 1.3) results from the volume and price effects described in section 2.1. Short-term oil revenues are higher under the Low Deployment scenario than in the Market Flooding scenario, but the situation is reversed after Peak Oil. In both scenarios, the post-Peak Oil rise of oil prices induces a surge of oil revenues; this surge is amplified in the Market Flooding scenario because of higher long-term oil prices. In this scenario, Middle-East countries can thus expect a reward for sacrificing short-term revenues and the trade-off between these two strategies depends on the objective function of Middle-East countries. Let us consider two polar objective functions as extreme cases where they put all weight on private interests (by maximizing oil revenues) or on the public welfare (by maximizing domestic households' surplus).

Figure 1.3. *Middle-East annual oil profits (Billion$)*



In the first case, Middle-East oil companies act as profit maximizing firms independent from any political influence. They choose their strategy based on discounted cumulated oil revenues (Table 1.2) and adopt the Market Flooding option only for discount rates lower than 6%. This is far below the high internal rates of returns demanded by private oil companies (17.26% to 21.97%, according to the Texas Comptroller's Property Tax Division[6]). Even though the recent financial crisis casts doubts upon the persistence of so high a profitability ratio, a breakeven point as low as 6% suggests that the adoption of the Market Flooding scenario is unlikely under this decision criterion (see Adelman (1986) for a more detailed analysis of discounting in the specific case of major oil producing countries).

[6] Determination of 2002 Discount Rate Range for Petroleum and Hard Mineral (*available at*: *http://www.window.state.tx.us/taxinfo/proptax/drs02/)*

Table 1.2. *Middle-East's discounted oil profits (Billion $)*

Discount rate[*]	*Limited Deployment Scenario*	*Market Flooding Scenario*
0%	38.9	**43.6**
1%	28.9	**31.8**
2%	21.9	**23.6**
5%	10.6	**10.8**
6%	**8.7**	8.6
7%	**7.2**	7.0
15%	**2.4**	2.2

[*]We present results for a selection of discount rates around the threshold values 5-6% defining the range of interest for the analysis,

Let us now assume that Middle-Eastern companies are managed in function of long-term public objectives. This means that Middle-East countries impose upon oil companies and sovereign funds to adopt pricing and investment decisions that maximize their households' surplus and to compare the general equilibrium effects of the two pricing strategies. Table 1.3 reports the variation of the population's surplus ΔS between the two scenarios: $\Delta S = \Delta R - CVI$, where ΔR and CVI are the effective and compensative variation of income respectively, the latter measuring the amount of income that would leave utility unchanged, given changes in relative prices. With this criterion, the Market Flooding scenario becomes a workable alternative because the social discount rate is lower than the private one, and because the range of discount rates for which the Market Flooding scenario is desirable proves to be much wider than with the oil profit maximization criterion: $\left[0\% - 13\%\right]$ instead of $\left[0\% - 7\%\right]$

Table 1.3. *Difference in Households' Surplus in the LD scenario with respect to the MF scenario (Billion $)*

Discount rate[*]	*Discounted surplus in LD w.r.t. MF*
5%	-1862
10%	-251
13%	-30
14%	+3
15%	+26
20%	+58

[*]We present results for a selection of discount rates around the threshold values 13-14% defining the range of interest for the analysis,

The difference between the two results originates in the long term macroeconomic effects of the two investment strategies. For a given assumption about the balance of payments, high short-term oil export revenues in the Low Deployment scenario are consistent with higher imports of industrial goods and a higher exchange rate of local currencies. This penalizes local industry and slows the transition of Middle-East countries away from oil-based revenues towards industrialization. Conversely, in the Market Flooding scenario, lower oil revenues allow for lower exchange rates. The development of local industry partially offsets short-term losses in oil revenues and better prepares Middle-East countries for the post oil era. Short-term inflows of oil revenues come at a pace compatible with the absorption capacity of the local economy, and the high post-Peak Oil inflows benefit to a more mature industrial structure. This captures in a simple form the 'natural resource curse' (Sachs and Warner, 2001) and the 'Dutch Disease': high resource rents do not guarantee sustainable growth patterns if limits in the absorption capacity of the economy weaken efficient re-investment in non-rent production sectors.

3- The adverse effects of cheap oil in oil-importing countries

Over the 2010-2050 time period, average GDP growth rates in the OECD are estimated to be 1.57% in the Low Deployment scenario *vs.* 1.53% in the Market Flooding scenario. These differences appear to be small in terms of discounted consumption (0.92% with a 2% pure time preference) or when translated into a growth delay (13 months). However, these aggregate indicators hide more significant discrepancies in the time profile of economic growth in OECD (Table 1.4).

Table 1.4. *Average growth rates in OECD (%)*

		Total (2010-2050)	Short-term Period (2010-2025)	Peak Oil Period (2025-2040)	Long-term Period (2040-2050)
Natural growth rates		1.42%	1.69%	1.30%	1.19%
Effective growth rates	Limited Deployment scenario	1.57%	1.93%	1.43%	1.24%
	Market Flooding scenario	1.53%	2.00%	1.29%	1.18%

An interesting indicator to investigate the importance of these time dependencies is the difference between natural and effective growth at different time horizons. Indeed, when effective growth is lower than (or very close to) the natural rate, it is impossible to avoid tensions in sectors or regions that are below this average effective growth, and to absorb the total labor force at constant wages. This happens in particular when investment and technical constraints inhibit the reallocation of the labor force towards the more productive sectors. Table 4 shows that the effective growth exceeds natural growth over the whole “pre-Peak Oil period” in the Market Flooding scenario and logically allows for higher OECD growth rates due to cheaper oil imports and cheaper energy for households and enterprises. During the “Peak Oil period”, the slowing down of economic growth starts sooner in the Market Flooding scenario and is more intense because Peak Oil hurts a more oil-dependent economy. During that period, the effective growth rate falls below the natural one for 10 years (2030-2040) in the Market Flooding scenario and continues to do so between 2040 and 2047. This corresponds to periods with high risks of social tensions. This situation never happens in the Low Deployment scenario.

These results lead to the conclusion that low energy prices over the short term are not necessarily beneficial for oil-importing countries since they may trap them in an oil dependency causing a strong variability of economic activity and lasting economic stagnation around and after the Peak Oil.

III. Uncertainties and their economic implications

After focusing on median assumptions for major determinants of oil markets, let us now conduct a sensitivity analysis to show the linkages between the main economic indicators and alternative assumptions on:

- the regional and total amount of oil resources; given controversies between pessimistic and optimistic views about these resources, we test a number of alternative scenarios in which the amount of resources is a weighted average between two extremes: 3.5 Trillion (10^{12}) bbl as a higher bound (2.3 Trillion bbl remaining conventional and 1.2 Trillion bbl of non conventional resource, in line with IEA (2008) estimates which gives a range for non conventional resources from 1 to 2Trillion bbl) and 2.4 Trillion bbl as a lower bound (1.6 Trillion bbl conventional and 0.8 Trillion bbl of non conventional), in line with estimates from the Association for the Study of Peak Oil (ASPO). The weighting factor m takes the value 0 and 1 for the lower and higher bounds respectively, and 0.5 in the central scenario analyzed in section II.
- the inertias affecting the deployment of non-conventional production; we consider four values of the parameter b to represent uncertainties on the rate of deployment of non-conventional oil: 0.07; 0.06 (value used in section II); 0.05, and 0.04. A higher b-value means an easier exploitation and faster deployment of non-conventional resources.

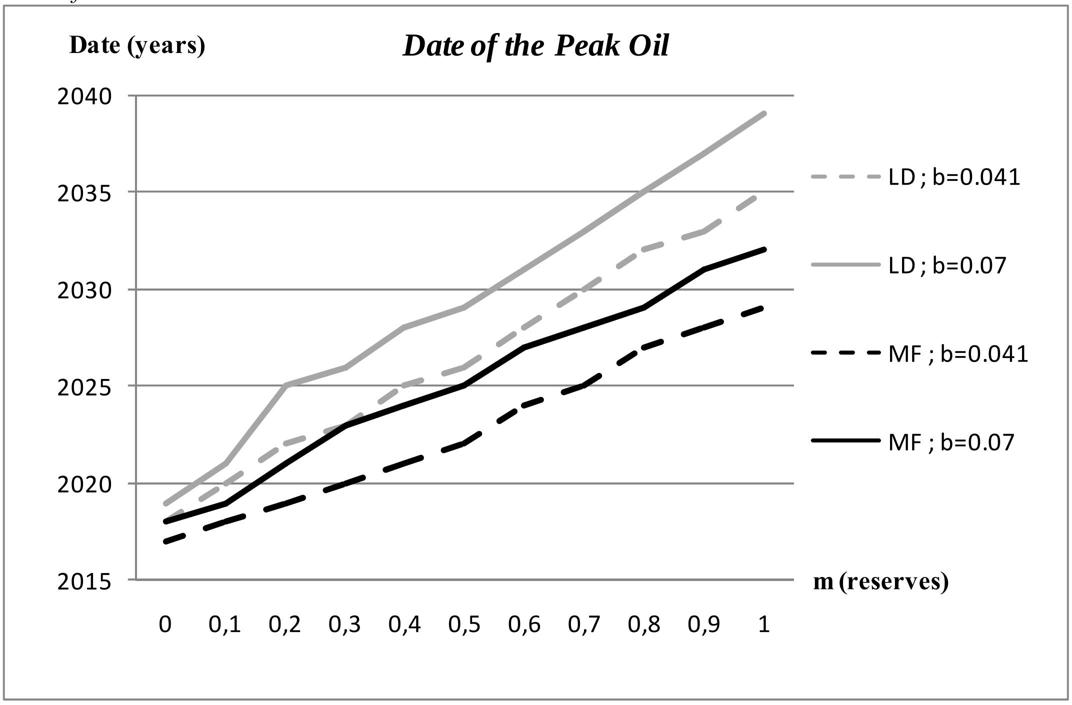
1- Early or late 'Peak Oil'? Geological uncertainties matter more that OPEC strategies

Figure 1.4 demonstrates a wide range of Peak Oil dates, from 2017 to 2039. Unsurprisingly, the size of the ultimate oil resource is the major determinant of this 22 year range, as shown by the strong increase of all curves from left to right. This figure also confirms the diagnosis of the median case analysis: for moderate assumptions on oil reserves, Peak Oil dates are weakly sensitive to the short-term price trajectory (the difference between Market Flooding and Low Deployment scenarios does not exceed five years for $m<0.5$).

In contrast, in case of abundant reserves, the sensitivity analysis demonstrates that the Peak Oil date depends significantly on other determinants. With $m=1$, the Peak Oil date varies by

11 years with respect to the selected pricing trajectory and technical parameters on non-conventional oil. This represents half the range of variations in Peak Oil dates and confirms a basic intuition of the paper, namely that, although the amount of reserves is an important factor, other economic and technical parameters may also play a key role in the determination of Peak Oil date.

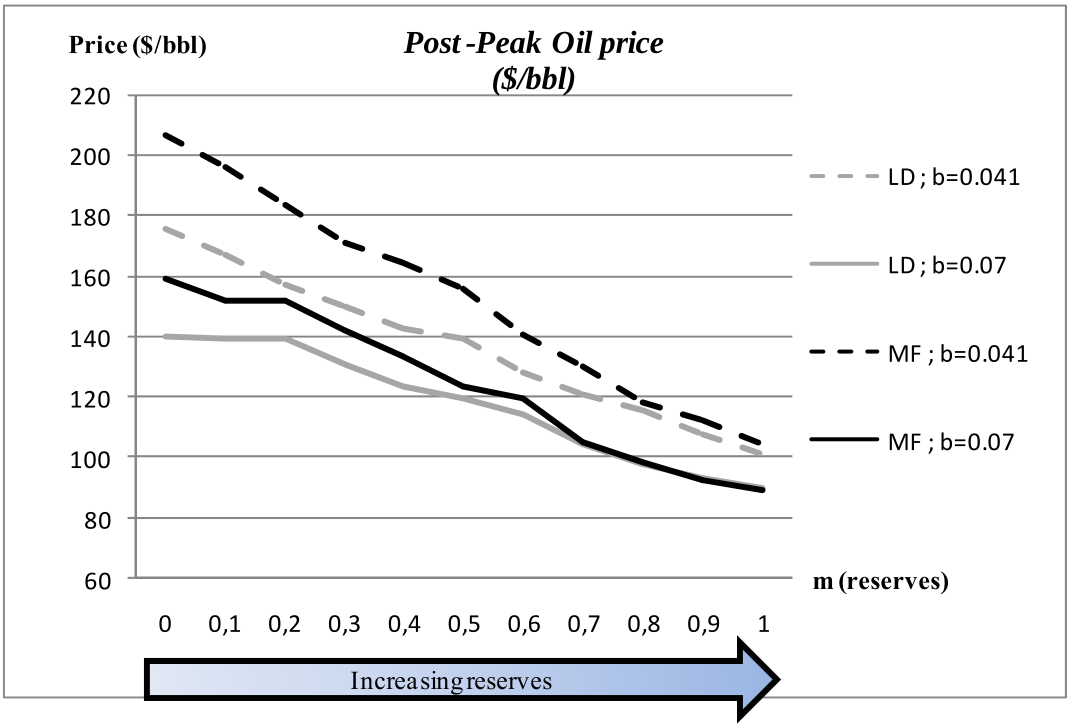
Figure 1.4. *Sensitivity of the date of Peak Oil with respect to the amount of resources and inertia in the deployment of non-conventional oil*



2- Long-term oil prices after Peak Oil

We now investigate the sensitivity of the average value of oil prices in the post-Peak Oil period, which is an indicator of tensions on oil markets (Figure 1.5). First, higher ultimate resources result in lower long-term oil prices as captured by the decreasing trend of all cruves from left to right. Indeed, *ceteris paribus*, higher resource gives a longer period for deploying oil-saving technologies and makes the economy less oil-dependent after Peak Oil. Second, long-term prices are always higher under a Market Flooding scenario because, misled by low price signals, oil-importing economies adopt more oil dependent consumption patterns triggering high demand. Third, optimistic views on non-conventional oil logically favor lower long-term prices by allowing a timely diffusion of substitutes to conventional oil, and hence helping to reduce the supply-side constraints on oil markets.

Figure 1.5. *Mean oil price during the post-Peak Oil period with respect to the amount of resources and inertia on the deployment of non-conventional oil*

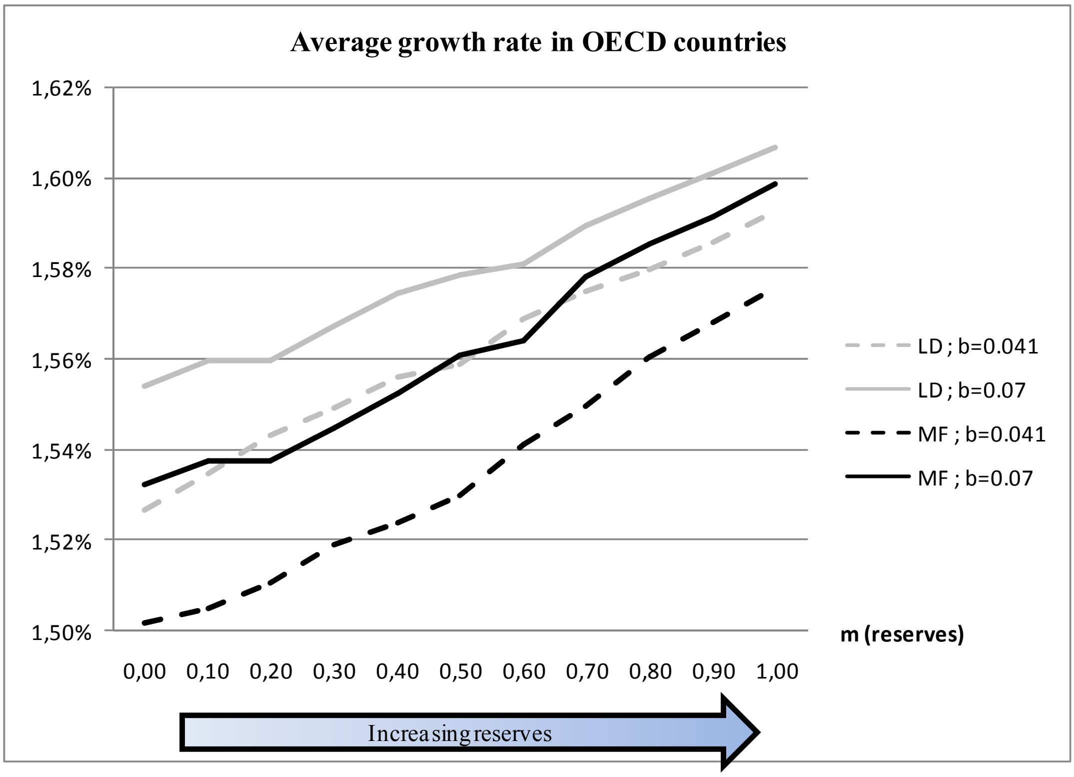


Interestingly, the comparison of sensitivity tests in 3.1 and 3.2 confirms that the date of Peak Oil says nothing about the time profile of oil prices. Indeed, in low resource cases, the date of Peak Oil is almost independent of parametric assumptions on pricing trajectories and inertias on the deployment of non-conventional oil, but these assumptions have a strong influence on long-term oil prices because they determine the abruptness of the break in demand and supply trends. Conversely, under high reserves, the wide range of Peak Oil dates hardly affects long-term oil prices, which remain moderate in all cases; indeed, Peak Oil happens late (not before 2028 under the more optimistic reserve assumption) so that oil-free technical change and the diffusion of substitutes to conventional oil have sufficiently progressed to limit the abruptness of the break in production and consumption trends at the Peak Oil period.

3- Macroeconomic effects and oil uncertainties

The analysis in Figure 1.6 shows that, unsurprisingly, more abundant reserves foster faster OECD growth by offering more abundant resource to these oil-importing economies. It also confirms for all parametric assumption that the Low Deployment scenarios are more profitable for OECD economies as they reduce their vulnerability to Peak Oil. On average this benefit is small and rather insensitive to parametric assumptions (less than 0.1% difference between the more extreme cases).

Figure 1.6. *Average growth rate in OECD countries with respect to the amount of resources and inertia on the deployment of non-conventional oil*



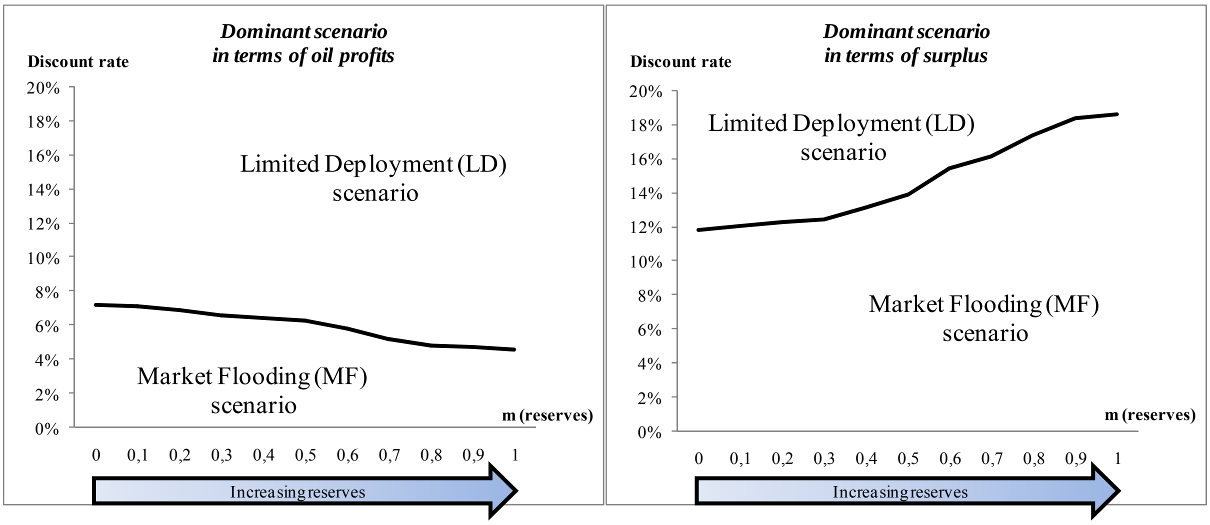
However, like in the median case, a much more contrasted picture is obtained when considering the time profiles. In particular, with low reserves, strong inertias on the deployment of non conventional oil and low short term oil prices, economic growth remains quite below the natural growth rate during 25 years after Peak Oil (Table 1.5), which is indicative of long lasting economic tensions.

Table 1.5. *Sensitivity tests on the time profile of OECD growth rates (%)*

		Short-term Period (2010-2025)	Peak Oil Period (2025-2040)	Long-term Period (2040-2050)
Natural growth rates		1.69%	1.30%	1.19%
Effective growth rates	Minimum	1.85%	1.19%	1.10%
	Maximum	2.05%	1.48%	1.26%

The situation is different for oil exporters, which appear more sensitive to parametric assumptions even for aggregate indicators like discounted revenues and discounted economic activity. We analyze these effects by delineating the domains of discount rates and resources over which each pricing scenario is dominant for Middle-East producers under the two decision criteria described in section 2.2 (figure 1.7).

Figure 1.7. *Dominant scenario for Middle-East countries with respect to the amount of resources and discount rate in terms of oil profits (left panel) and surplus (right panel).*



In all scenarios, higher resources decrease discounted Middle-East oil revenues, since later Peak Oil postpones the bubble of long-term oil revenues and limits its magnitude due to lower long-term oil prices. The magnitude of this effect depends on the scenario considered whilst the amount of reserves also infleunces Middle-East producers' trade-off between the MF and Low Deployment scenarios.

When producers act as private companies, the threshold value for discount rates remains low (5-7%) and the trade-off favors the Low Deployment scenario for all assumptions (Figure 7, left panel). When considering social surplus, threshold discount rates are much higher and delineate a notably wider dominant domain for the Market Flooding scenario (Figure 7, right panel).

More remarkably, for economically meaningful reasons, the trend of the curves with respect to the amount of resources depends on the decision criterion. The downward oriented slope in (Figure 7, left panel) demonstrates that the Market Flooding scenario is penalized by high resources with private assessments. Indeed, higher resources lead to a longer period of technical change before constraints on oil supply appear, and oil-importing economies are less oil-dependent when hit by ‘Peak Oil’. This leads to a delayed and lower long-term bubble of oil profits which affects the reward for the short-term sacrifice.

When considering social assessments instead, the upward oriented slope in (Figure 7, right panel) demonstrates that the Market Flooding scenario is favored by high resources. This is due to the impact of oil resources on the magnitude and duration of the *Dutch Disease* mechanism and on the length of the period during which oil importers are directed towards oil-intensive pathways. Higher resources extend the period during which lower oil revenues in the Market Flooding scenario force the development of local industrial production in Middle-East countries. In this way, the long-run absorption capacity of Middle-East economies is improved after Peak Oil, i.e. at the moment when they get the bubble of oil revenues.

IV. Conclusion

This paper reviews the notion of Peak Oil in a general equilibrium modeling framework that represents the limits on the short term adaptability of oil supply, oil substitutes and fuel demand. In this framework, inertia and imperfect foresight create the possibility of a sudden acceleration in oil price increases if importing economies are very oil-dependent when entering the period of oil depletion.

By considering two counterfactual scenarios, sensitivity tests show that the date of Peak Oil is sensitive to short-term oil price only in case of high reserves and that Peak Oil dates that differ only slightly may lead to very different time profiles of oil prices, rent formation and growth patterns.

From oil exporters’ point of view, low oil prices undermine short-term exportation revenues; but they encourage oil consumption, make oil-importing economies more oil-dependent at the Peak Oil date and create room for a bubble of long-term oil exportation revenues. It thus may be in the interests of oil producers to accept a temporary sacrifice in their short-term export revenues so as to benefit from higher long-term revenues in the post-Peak Oil period. But, they will do so only if they consider long-term macroeconomic objectives (including industrialization and hedging against Dutch Disease) instead of the maximization of discounted oil revenues. This option is all the more attractive in case of high reserves.

From oil importers' point of view, long periods of low energy prices make the economy more vulnerable to Peak Oil and may not ultimately be beneficial. It may thus be in their interest to correct potentially misleading price-signals by using complementary measures to secure steady technical change. Among them, international climate policies can be envisaged as a hedging strategy against the negative long-term economic outcome of the uncertainty on oil markets (Rozenberg et al, 2010). This possibility, in turn, raises the question of Middle-East countries' reaction to these measures.

REFERENCES

Adelman M A (1986). Oil producing countries discount rates. Resources and Energy 8(4): 309-329.

Ahmad S (1966). On the theory of induced innovation. Economic Journal 76: 344-357.

Al-Husseini M (2006). The debate over Hubbert's Peak: a review. GeoArabia 11(2): 181:210

Amigues JP, Favard P, Gaudet G, Moreaux M (1998). On the Optimal Order of Natural Resource Use When the Capacity of the Inexhaustible Substitute Is Limited. Journal of Economic Theory 80(1):153-170

Anderson KP (1972). Optimal Growth When the Stock of Resources is finite and depletable. Journal of Economic Theory 4: 256-267.

Atkeson A, Kehoe PJ (1999). Models of energy use: putty-putty versus putty-clay. The American Economic Review 89(4): 1028-1043.

Barro RJ, Sala-i-Martin X (1992). Convergence. Journal of Political Economy 100(21): 223-251.

Bentley RW, Mannan S, Wheeler S (2007). Assessing the date of the global oil peak: The need to use 2P reserves. Energy Policy 35(12): 6364-6382

Bresnahan TF, Ramey VA (1993). Segment Shifts and CapacityUtilization in the U.S. Automobile Industry. American Economic Review Papers and Proceedings 83(2): 213-218.

Corrado C, Mattey J (1997). Capacity Utilization. Journal of Economic Perspectives, 11(1): 151-167.

Davis S, Haltiwanger J (2001). Sectoral Job Creation and Destruction Responses to Oil Price Changes. Journal of Monetary Economics 48: 465-512.

Deffeyes KS (2002). Hubbert's Peak: The Impending World Oil Shortage. Princeton, New Jersey, Princeton University Press.

Dimaranan BV (2006). Global Trade, Assistance, and Production: The GTAP 6 Data Base. Center for Global Trade Analysis, Purdue University.

Fattouh B (2007). The drivers of oil prices: the usefulness and limitations of non-structural model, the demand-supply framework and informal approaches. Working Paper No. WPM32, Oxford Institute for Energy Studies.

Finn MG (2000). Perfect competition and the effects of energy price increases on economic activity. Journal of Money, Credit and Banking 32(3): 400-416.

Frondel M, Schmidt C (2002). The Capital-Energy Controversy: an artifact of cost shares? The Energy Journal 23(3): 53-80.

Greene DL, Hopson JL, Li J (2006). Have we run out of oil yet? Oil peaking analysis from an optimist's perspective. Energy Policy 34(5): 515-531.

Greening L, Greene DL, Difiglio C (2000). Energy efficiency and consumption – the rebound effect – a survey. Energy Policy 28 (6-7): 389-401.

Guivarch C, Hallegatte S, Crassous R (2009). The resilience of the Indian economy to rising oil prices as a validation test for a global energy–environment–economy CGE model. Energy Policy 37(11): 4259–4266.

Hamilton JD (2008). Oil and the Macroeconomy. The New Palgrave Dictionary of Economics.

Holland S (2008). Modeling Peak Oil. The Energy Journal 29(2): 61-80.

Holland S (2003). Extraction Capacity and the Optimal Order of Extraction. Journal of Environmental Economics and Management 45(3): 569-588.

Hotelling H (1931). The Economics of Exhaustible Resources. Journal of Political Economy 39: 137-175.

Hourcade JC, Jaccard M., Bataille C., Ghersi F (2006). Hybrid Modeling: New Answers to Old Challenges**.** The Energy Journal, Special Issue n°2: 1-11.

Hubbert MK (1962). Energy Resources. A Report to the Committee on Natural Resources. National Academy of Science. Government Printing Office, Publication No. 1000-D.

Hubbert MK (1956). Nuclear Energy and the Fossil Fuels. American Petroleum Institute Drilling and Production Practice, Proceedings of Spring Meeting, San Antonio, 7-25.

International Energy Agency (IEA), 2008. World Energy Outlook 2008. IEA/OCDE, Paris.

International Energy Agency (IEA), 2006. Energy Technology Perspectives: Scenarios and Strategies to 2050. OECD/IEA, Paris, France.

Intergovernmental Panel on Climate Change (IPCC), 2007. Summary for Policymakers. In: B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds) Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. CambridgeUniversity Press, Cambridge, United Kingdom and New York, NY, USA.

Kaufmann RK, Dees S, Karadeloglou P, Sanchez M (2004). Does OPEC matter? An econometric analysis of oil prices. The Energy Journal 25(4): 67-90.

Kemp MC and Van Long N (1980). On Two Folk Theorems Concerning the Extraction of Exhaustible Resources. Econometrica 48 (3): 663-673

Krautkraemer JA (1998). Nonrenewable Resource Scarcity. Journal of Economic Literature 36: 2065-2107.

Lynch MC (2003). The New Pessimism about Petroleum Resources: Debunking the 'Hubbert Model' (and Hubbert Modelers). Minerals and Energy 18(1): 21-32.

Maddison A (1995). Monitoring the World Economy:1820 – 1992, OECD Development Center, 260 p.

Oliveira Martins J, Gonand F, Antolin P, de la Maisonneuve C (2005). The impact of ageing on demand, factor markets and growth, *OECD Economics Department Working Papers*, #420, OECD Economics Department.

Phelps E (1961). The Golden Rule of Accumulation: A Fable for Growthmen. The American Economic Review 51(4): 638-643

Rehrl T, Friedrich R (2006). Modeling long-term oil price and extraction with a Hubbert approach: The LOPEX model. Energy Policy 34(15): 2413-2428.

Reynolds DB (1999). The mineral economy: how prices and costs can falsely signal decreasing scarcity. Ecological Economics 31 (1): 155–166.

Rogner H-H (1997). An assessment of world hydrocarbon resources". Annual Review of Energy and the Environment 22: 217–262.

Rotemberg J, Woodford M (1996). Imperfect Competition and the Effects of Energy Price Increases. Journal of Money, Credit, and Banking 28: 549-577.

Rozenberg J, Hallegatte S, Vogt-Schilb A, Sassi O, Guivarch C, Waisman H, Hourcade JC (2010). Climate policies as a hedge against the uncertainty on future oil supply. Climatic Change 101(3-4): 663-668

Sachs JD, Warner AM (2001). The Curse of Natural Resources. European Economic Review 45(4–6): 827–838.

Sands R., Miller S., Kim M-K. (2005).The Second Generation Model: Comparison of SGM and GTAP Approaches to Data Development", *Pacific Northwest National Laboratory*, PNNL-15467.

Schäfer A, Victor DG (2000). The future mobility of future population.Transportation Research Part A 34: 171-205.

Solow RM (1974). Intergenerational Equity and Exhaustible Resources. Review of Economic Studies (Symposium 1974): 29-45.

Stiglitz J (1974). Growth with Exhaustible Natural Resources. Review of Economic Studies (Symposium 1974): 123-152

Uhler RS (1976). Costs and supply in petroleum exploration: the case of Alberta. Canadian Journal of Economics 19: 72–90.

United States Geological Survey (USGS), 2000. World Petroleum Assessment 2000. USGS, Washington.

Waisman H, Guivarch C, Grazi F, Hourcade JC (2011). The Imaclim-R Model: Infrastructures, Technical Inertia and the Costs of Low Carbon Futures under Imperfect Foresight, *In:* The Economics of Ambitious GHG Emission Reductions, Special Issue, Climatic Change, *accepted for publication.*

Chapter 2

Climate policy and oil exporters: a general equilibrium assessment of monetary compensations

This second chapter* extends the modeling exercise of Chapter 1 to consider the effect of a climate policy on oil markets and the macroeconomy. This allows us to disentangle the interactions between a worldwide carbon price, oil prices and production patterns, and to quantify the consequences on Middle-East producers' economy in terms of exportation and macroeconomic losses. This assessment gives the magnitude of monetary transfers oil producers may claim to compensate for their losses, an often discussed point in climate negotiations. The evaluation of macroeconomic effects on both the receivers and the contributors of these transfers proves that, because of its magnitude and time profile, the compensation of GDP losses is both more acceptable and efficient than the compensation of oil exportation revenue losses. We finally consider the case where these compensations are not Middle-East countries do not participate to the climate coalition. By comparing the policy costs of the additional carbon burden with those due to monetary compensations under two oil pricing trajectories, we delineate the geopolitical interactions between oil markets and climate negotiations.

* This chapter is the outcome of a joint collaboration with Jean-Charles Hourcade and Julie Rozenberg.

From its introduction on the diplomatic agenda in 1988, the climate change issue has been closely related to energy security concerns. The co-benefit of climate policies in terms of reduced dependency on fossil fuels' importations (Schlessinger, 1989) has remained a major political justification for the adoption of ambitious post-2012 climate targets in oil-importing countries even during the periods of low oil prices.[1] The corollary is the concern of the Organization of Petroleum-Exporting Countries (OPEC) about the resulting drop of their exportation revenues. This adverse impact of the climate policy has led oil exporters to claim for monetary compensations in the context of climate negotiations since the UNFCCC (Article 4.8) and the Kyoto protocol (article 3.14) (Barnett and Dessai, 2002). Although it conditions to a great extent the participation of Middle-East countries to a climate policy[2], this question has remained a major unadressed political issue until now.[3]

Beyond purely geopolitical obstacles, the difficulty to reach an agreement on this question refers also to the controversial picture given by the economic literature on the socio-economic consequences of climate policy in oil producing countries. On the one hand, hybrid energy-economy models concur in predicting significant losses for oil exporters at both short and long term[4]. These analyses are suited to capture the general equilibrium feedbacks causing adverse impacts, but they conventially do so with first-best assumptions minimizing the possibility of co-benefits with respect to the optimal baseline. On the other hand, dynamic partial equilibrium models obtain that ambitious climate targets may benefit to conventional

[1] The Gleneagles Communiqué following the G8 summit in 2005 is a symptomatic example of this close interplay: "(a) Climate change is a serious and long-term challenge […] We know that increased need and use of energy from fossil fuels contribute in large part to increases in greenhouse gases associated with the warming of our Earth's surface; (b) Global energy demands are expected to grow by 25% over the next 25 years. This has the potential to cause a significant increase in greenhouse gas emissions associated with climate change […]; (c) Secure, reliable and affordable energy sources are fundamental to economic stability and development. Rising energy demand poses a challenge to energy security given increased reliance on global energy markets.[…]"

[2] as illustrated by the statement of Dr. Rilwanu Lukman, the Secretary General of OPEC at COP-4: 'without a favourable disposition towards the compensation issue among the Parties to the Convention, how can fossil fuel producers be expected to give their wholehearted blessing to measures that could wreak havoc with their economies?'

[3] The 2009 Copenhagen Accord only recognizes "the potential impacts of response measures on countries particularly vulnerable to its adverse effects" but does not provide concrete advancement on this question.

[4] The full implementation of the Kyoto Protocol over a 10 year period reduces oil exportation revenues by 9.8% to 13% in 2010 with a few percentage point losses in terms of economic activity and welfare (Barnett et al., 2004). At a 2050 horizon, a 550ppm target would induce a 35% decrease of oil revenues in OPEC countries (Van Vuuren et al., 2003) and total abatement cost at approximately 2 percent of GDP in Middle East (WBGU, 2003).

oil producers[5]. These results reflect the supply constraints imposed by different fractions of oil (conventional oil, tar sand, heavy crude oil, and oil shale), but the partial equilibrium nature of the exercise fails to consider the feedback effects of climate policies on oil demand through reduced economic activity, a major adverse impact on Middle-East economies.

This paper bridges the gap between these two branches of the literature by representing different categories of oil in a Computable General Equilibrim (CGE) energy-economy model. But, we also extend these conventional approaches to represent the transitory departures from first best trajectories due to limited flexibility of economic adjustments. The basic intuition is indeed that, under imperfect expectations, adverse impacts and co-benefits of a climate policy depend crucially upon market imperfections (in particular, OPEC's market power) and technical inertias on *(i)* the deployment of oil production capacities, *(ii)* the diffusion of alternatives to oil, *(iii)* the capital turnover in infrastructure sectors, and *(iv)* the pace and direction of end-use technical change.

Section 1 sketches the rationale of Middle-East producers' decisions in terms of oil pricing stategy and compliance to a climate agreement with and without compensations. Section 2 describes and discusses the modeling assumptions behind the representation of market imperfections and technical inertias in a CGE framework. With this framework, Section 3 estimates the socio-economic consequences of climate policy, including its adverse impacts on Middle-East countries and co-benefits for energy importers. This assessment serves in Section 4 as a basis for estimating the monetary compensations Middle-East producers may claim to oil importing countries, and for discussing their acceptability regarding the amount of associated transfers and the economic consequences for contributors. Section 5 considers finally the possibility that Middle-East exits the coalition and examines the additional burden it imposes on other regions.

[5] by affecting more the cost of their potential substitutes (unconventional oil, coal) (Persson et al., 2007) and fostering an increase of conventional oil rents in OPEC if they can exert their market power (Johansson et al, 2009)

I. OPEC's decisions in a climate policy context

Faced with a context of global agreement on climate policies, major oil producers can envisage three types of attitudes vis-à-vis the coalition:

- participating to the global agreement without compensation (section III). This corresponds to an ideal case which may appear quite unrealistic after Copenhagen and Cancun, but constitutes a useful benchmark to evaluate the ultimate impacts of the climate policy.
- participating to the climate coalition in exchange of a compensation for the adverse effects of the climate policy (section IV). In this case, OECD countries agree to compensate the adverse impacts of the climate policy for oil producers through monetary transfers.
- remaining outside of the global agreement (section V). This option imposes an additional burden to remaining participants to the climate coalition if they want to achieve the same climate target. Its rationale for Middle-East countries is to avoid the slowing down of economic development induced by a local carbon constraint, but also to try and discourage the implementation of climate policies by undermining their political acceptability through enhanced costs in OECD countries.

Although depending on a complex interplay of domestic and geopolitical determinants that are far beyond the scope of this paper, the political acceptability of either option driving the selection of one of these attitudes is largely dependent upon their ultimate economic consequences. These consequences in turn depend upon the way Middle-East countries will use the other strategic leeway at their disposal, which is the ability to influence world oil prices.

The way Middle-East countries will exert their market power is conditional upon their ability to elaborate coordinated strategies during a temporary period. Although the climate policy may intensify intra-cartel tensions due to its lowering effect on oil revenues[6], we admit that it will not affect the cartel discipline. This assumption is in line with (Berg et al, 1997a) who show that gains from cartelization are not that much impacted by global carbon taxes.

[6]This is particularly true since there are great discrepancies among Middle-East countries in terms of oil revenues per capita (more than 5000$ per capita in Saudi Arabia *versus* 800$ in Iran)

The response of this cartel to various profiles of carbon prices have been widely investigated (IPCC, 2001, section 8.3.2.3) and the conclusion is often that OPEC would adopt a limited deployment of oil production capacities to cut back production and maintain revenues (Berg and al., 1997b). However, given the experience of the mid-eighties, a market flooding strategy with low short-term oil prices cannot be a priori excluded. [7] This is why we will consider two polar options for short-term oil pricing strategies (note that in the IMACLIM-R model, before depletion starts Middle-East producers can control the time profile of oil prices through the utilization rate of production capacities):

- *The Limited Deployment scenario (LD)*: Middle-East producers refrain from investing in new capacities and maintain the medium term oil price around p_{high}=90$/bl [8].
- *The Market Flooding scenario (MF):* the expansion of production capacities in Middle-East countries brings back oil price at its pre-2004 level, p_{low} = 50$/bl.

The Limited Deployment (LD) scenario permits to extract oil rents as soon as possible before the climate policy reduces significantly oil demand, contrary to the Market Flooding (MF) strategy in which short-term oil revenues remain moderate. This sacrifice may be made in view of two benefits. On the one hand, a low oil price makes a higher carbon price necessary to meet a given climate target. Since consumers and tax-payers are particularly sensitive to the carbon price level, this may undermine the political acceptability of climate policies in oil-importing countries and hence reduce the adverse impacts in Middle-East countries. On the other hand, under non-perfect expectations, low oil prices may lead oil importing countries towards oil intensive pathways.[9] Taking into account inertias on the deployment of oil production capacities, oil substitutes and demand-side technical change in OECD countries, the demand for oil remains important even in the long run and high scarcity rents can then be formed. In this pricing trajectory, low short-term revenues may thus be compensated by long-term profits.

[7] the recent ups and downs in oil prices down to 47$/barrel in November 2008 demonstrates the possibility of low prices for reasons other than an explicit market flooding strategy. However, the fall of prices after economic crisis simply demonstrates that OPEC did not cut back immediately production capacity.

[8] All price levels are expressed in $2001, the values of p_{low} and p_{high} corresponding respectively to around 50$/bl and 100 $/bl in current currency. The price level estimated by the Short-Term Energy Outlook 2010 (available at: http://www.eia.doe.gov/emeu/steo/pub/contents.html) is 78$/bl, and falls in the middle of the range defined by p_{low} and p_{high}.

[9] Such a strategy would replicate the drop of oil prices from 35$ to 8$/bbl between 1982 and 1985, followed by a stabilization at a rather moderate level, which weakened the discipline adopted in OECD countries in response to the oil shocks.

But, the time profile of oil revenues is only one indicator of costs and benefits of either strategy. Middle-East governments have indeed to consider broader objectives such as calming down social tensions or ensuring a sustainable long-run development beyond the 'oil era'. This is why we have to capture the induced socioeconomic effects as measured by macroeconomic indicators of economic activity and welfare through a general equilibrium analysis.

II The model

We adopt the energy-economy model IMACLIM-R, which has been designed to endogenize the interplay between the long-term dynamics of oil markets, the strategies of major oil producers and general equilibrium effects under climate policy. This model of the world economy[10] covers the period 2001-2050 in yearly steps through the recursive succession of annual static equilibria and dynamic modules. The *annual static equilibrium* determines relative prices, wages, labour, value, physical flows, capacity utilization, profit rates and savings at date *t* as a result of short term equilibrium conditions between demand and supply on goods, capital and labor markets. The *dynamic modules* are sector-specific reduced forms of technology-rich models, which take the static equilibria at *t* as an input, assess the reaction of technical systems to the economic signals, and send new input-output coefficients back to the static model to compute the equilibrium at *t*+1.

The detailed rationale of the IMACLIM-R model as well as the full set of equations of the static equilibrium and the description of all dynamic modules are provided in Annex A; we limit here to summarize the crucial modeling features adopted for the representation of the core mechanisms in the present paper, namely *(i)* oil supply constraints, *(ii)* general equilibrium effects of oil price variations and *(iii)* monetary compensations in the form of international capital transfers.

(i) The dynamic module describing oil supply determinants embarks three crucial specificities of oil supply (see Chapter 1 for a more detailed description). First, oil reserves are heterogeneous, which is captured by distinguishing different categories according to their

[10] The IMACLIM-R model used in this paper divides the economy in 12 regions—USA, Canada, Europe, OECD Pacific, Former Soviet Union, China, India, Brazil, Middle East, Africa, Rest of Asia, Rest of Latin America—, and 12 productive sectors—Coal, Crude Oil, Natural Gas, Refined products, Electricity, Construction, Agriculture and related industries, Energy-intensive Industries, Air Transport, Sea Transport, Other Transports, Other industries and Services. In addition IMACLIM-R includes transportation with personal vehicles and non-motorized transport.

nature (conventional *vs* non conventional) and their cost of exploration/exploitation; second, the geological nature of oil reserves imposes a limited adaptability of oil supply, which is captured by imposing a maximum annual rate of increase for production capacities of each oil category; third, a small group of suppliers (Middle-East countries) benefits from a market power, which allows them to influence world oil prices through their production decisions.

(ii) The magnitude of oil-macroeconomy interactions are reproduced thanks to the five following features, which have been pointed by (Hamilton, 2008) as crucial determinants of the economic effects of oil price variations (see Chapter 1 for a more detailed discussion) *1) mark-up pricing* to capture market imperfections (Rotemberg and Woodford, 1996); *2) partial utilization rate of capital* when the full utilization of installed production capacities cannot be secured due to limits in the substitution between capital and energy (Finn, 2000); *3) a putty-clay description of technologies* to represent the inertias in the renewal of capital stock (Atkeson and Kehoe, 1999); *4) frictions in the reallocation of capital across heterogeneous sectors* causing differentiated levels of idle production capacities (Bresnahan and Ramey, 1993); *5) frictions in the reallocation of labor across heterogeneous sectors* causing differentiated levels of unemployment (Davis and Haltiwanger, 2001).

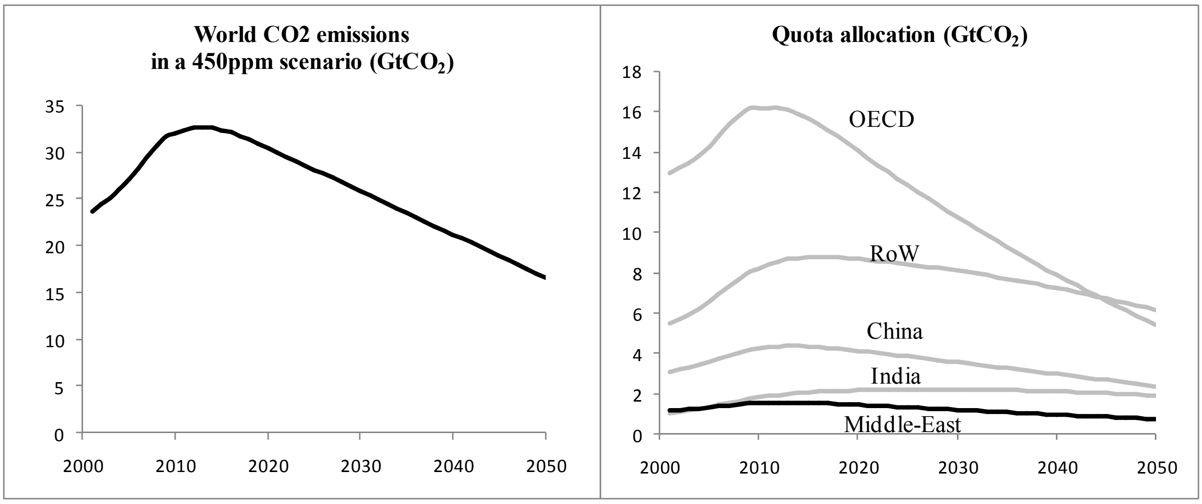
(iii) The general equilibrium effects of monetary transfers are captured through their impact on countries' current account and thus their trade balance. More specifically, the overall regional capital balance is given by the sum of these monetary transfers and the rest of capital flows, these latter being assumed to decrease exponentially over time as a progressive correction of international capital imbalances by 2050. Under the standard assumption of equilibrated balance of payments, these capital accounts (including monetary transfers) must be exactly compensated by trade flows on international markets[11]. This compensation happens through adjustments of the set of relative prices defining regional competitiveness, and ultimately affects each region's economy because it determines the terms of trade and hence the potentials for exports and the cost of imports.

[11] All goods are internationally tradable; for non-energy goods, Armington specifications (Armington, 1969) capture the partial substitutability between domestic and foreign goods, while physical accounting for energy goods (in MToe) makes them fully substitutable.

III. Costs of Climate policy for oil producers

In our modeling exercise, a climate policy is represented by a carbon tax associated to carbon emissions from the production and use of fossil energies (coal, oil and gas). It thus increases the cost of final goods and intermediate consumption according to the carbon content of the fuel used. At each date, the carbon price value is endogenously calculated to curve carbon emissions according to a prescribed objective. We consider a 450 ppm-CO_2 stabilization target, which imposes a peak of world CO_2 emissions between 2010 and 2020 and a decrease by 30% in 2050 with respect to 2000 levels (Barker et al., 2007, Table TS2) (Figure 2.1a). Associated revenues are collected by the government which then reallocates them to households and/or firms through transfers. An international permit market is modeled by introducing transfers according to the difference between real emissions and regional allocations, decided by a "Contraction and Convergence" principle (Figure 2.1b).

Figure 2.1. (a) *World CO_2 emissions under climate policy ($GtCO_2$) [left-hand panel]; (b) Regional quota allocation ($GtCO_2$) [right-hand panel].*

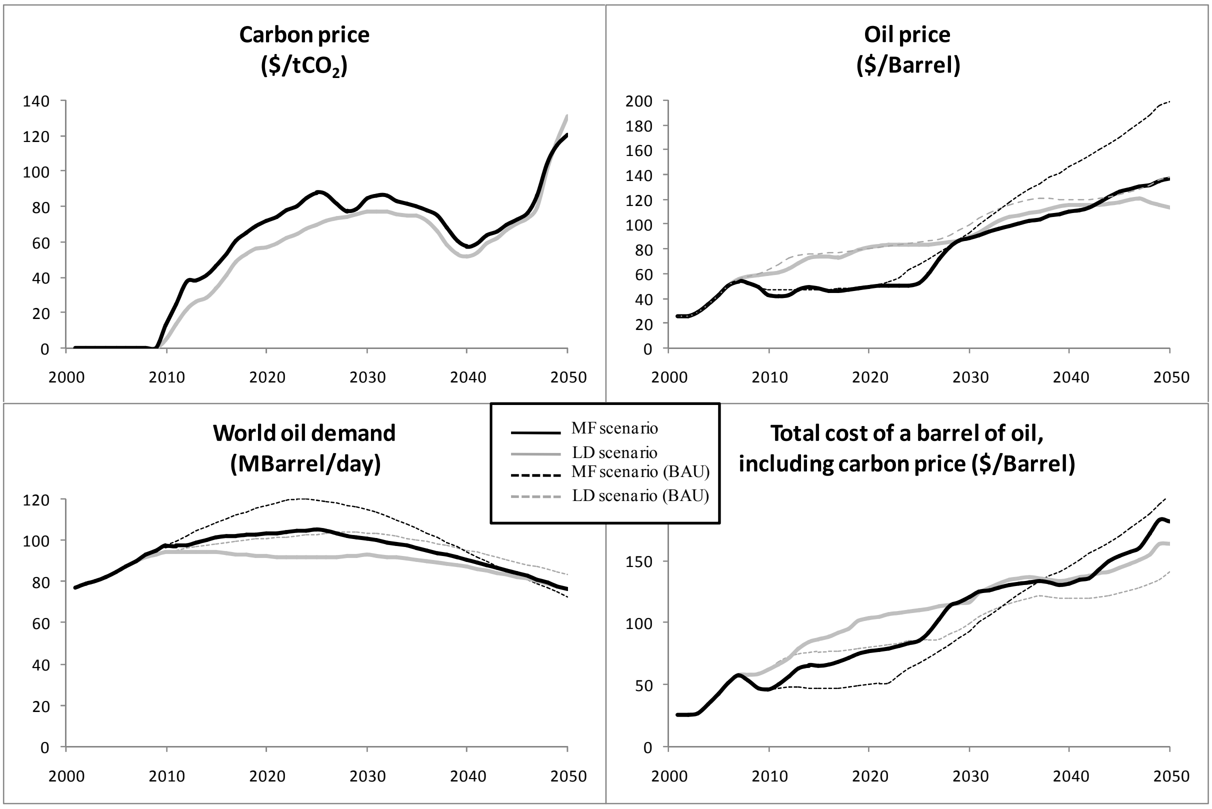


In this section, we consider the 'ideal' case of a world agreement on a Kyoto-type architecture with full participation. This assumption may seem unrealistic since it assumes that OPEC countries decide to accept carbon constraints without monetary compensations for their adverse impacts; it will however be used as a benchmark for sections IV and V.

1- Carbon pricing and oil markets

The carbon price follows a similar trajectory under the two oil pricing strategies, with three distinct periods (Figure 2.2a). During the first years of the climate policy (2010-2025), oil prices remain controlled by Middle-East producers and are then almost unaffected by the climate policy (Figure 2b). In this case, high carbon prices are necessary to ensure a steady increase of the cost of oil for final users (Figure 2d) in turn triggering the decrease of oil consumption necessary to comply with the reduction of carbon emissions (Figure 2c). Under the MF scenario, the increase of carbon prices is sensibly steeper (89\$/tCO_2 in 2025 *vs.* 70\$/tCO_2 under LD scenario), but total oil demand remains more important (105 MBarrel/day at the maximum vs. 97 MBarrel/day in the LD scenario) because of lower total cost of a barrel.

Figure 2.2. *(a) Carbon price (\$/tCO_2) [upper-left panel]; (b) Oil price (\$/Barrel) [upper-right panel]; (c) World Oil demand (MBarrel/day) [lower left panel]; (d) Total cost of a barrel of oil, including carbon costs (\$/Barrel) [lower-left panel]*



Between 2025 and 2040, the carbon price stagnates and even declines in both scenarios (Figure 2.2a). They remain above 50\$/tCO_2 to sustain the continuous decrease of oil demand compatible with the climate policy, which proves sufficient to tap most mitigation potentials in power, residential and industrial sectors, which represent the core of emissions reductions at that medium-term stage of the climate policy (Barker et al., 2007, Figure TS27). Oil demand declines continuously (Figure 2.2c), because of the steady increase of end-use costs (Figure 2.2d) caused by the post Peak Oil rise of prices (Figure 2.2b). During this period, carbon prices remain lower under the LD scenario because higher cost of fossil fuels in the previous period has accelerated carbon-free technical change.

After 2040, a steep increase of carbon prices is experienced in both scenarios (from around 60 \$/tCO_2 in 2040 to 120\$/tCO_2 in 2050) to reach the high-cost mitigation potentials in the transportation sector, which become important above 100\$/tCO_2 (Barker et al., 2007, Figure TS27). Oil prices are lower than in the baseline scenario since carbon pricing accelerates fossil-free technical change and thus reduces the long-term oil dependency of the economy. This effect of carbon prices on oil prices is moderate in the LD scenario, while particularly important in the MF scenario (Figure 2.2b). In this latter case, the lowering of oil prices under the climate policy is even sufficient to make oil products less expensive than in the baseline, even with the carbon price (Figure 2.2d), and hence to foster slightly higher oil demand than in the baseline in the last period (Figure 2.2c).

Logically, world oil demand under climate policy is always higher in the MF scenario than in the LD scenario over the whole period 2010-2050, with a significant cumulative difference amounting to 94 GBarrels (12.9 GToe). Table 2.1 shows a similar effect for gas as a result of the partial indexation of gas prices on low oil prices, with 19% higher cumulated gas consumption in the MF scenario over the 2010-2050 period. The compensation, to achieve identical climate objectives, will be made through lower coal-related emissions (-12%). This analysis demonstrates that maintaining low oil prices is an efficient way to protect the oil share in energy markets under a climate policy, by redirecting the mitigation efforts towards a more intense reduction of coal.

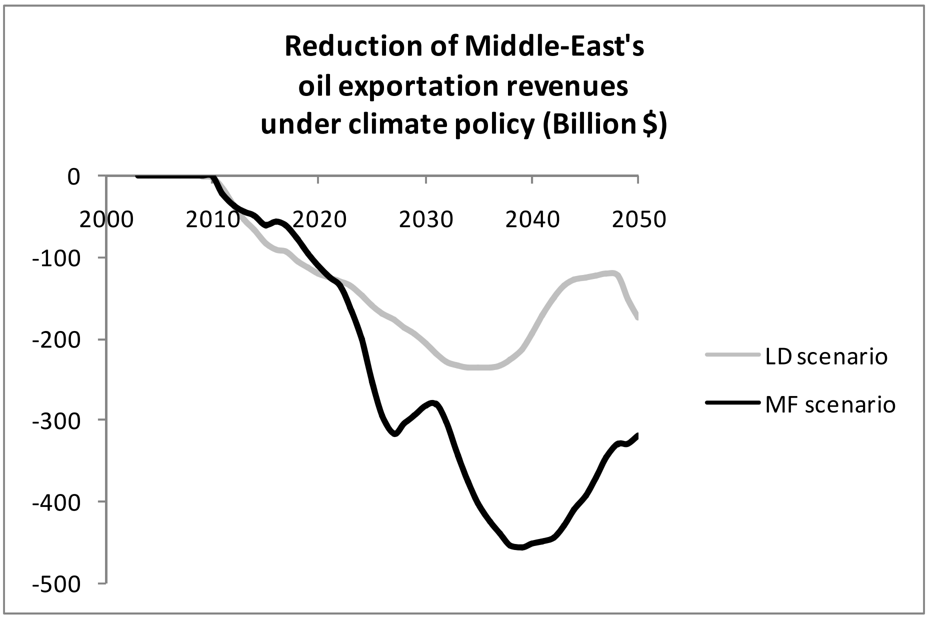
Table 2.1. *Cumulated production of fossil fuels under climate policy over the period 2010-2050 (GToe)*

	Oil	Gas	Coal
LD scenario	182.3	97	162.6
MF scenario	195.2	115.1	142.9

2- The climate policy and Middle-East countries' economy

Unsurprisingly, the climate policy induces a significant drop of oil revenues in Middle-East countries due to the combination of lower prices and demand (Figure 2.3). Between 2010 and 2020, the OPEC maintains prices at their baseline levels and the reduction of oil exportation revenues is due to the adverse impact on oil demand. Cumulative losses over this period are more important in the LD scenario in absolute terms (1032 Billion $ against 870 Billion $ in the MF scenario), and represent a similar reduction of total oil revenues over the period in the two scenarios (15.7% and 17.5% respectively). Over the whole period 2010-2050, the long run effects of climate policy on oil prices, especially in the MF scenario, enhance the losses, which reach 10 700 Billion$ losses, or 26% of cumulated oil revenues over. Indeed, the climate policy delays the Peak Oil and minors its consequences in terms of rising prices, hence strongly limiting the bubble of long-term profits Middle-East countries receive in the baseline case. This effect is also at play under the LD scenario but with a minored magnitude: cumulative losses are limited to 6 000 Billion$ losses over 2010-2050, or 16% of total oil revenues over the period.

Figure 2.3. *Reduction of oil exportation revenues under climate policy (Billion $)*



Less intuitive is the somewhat different picture obtained by observing the economic trajectories in Middle-East countries, as measured by losses of domestic economic activity in GDP terms when they comply to the climate coalition (Figure 2.4a).

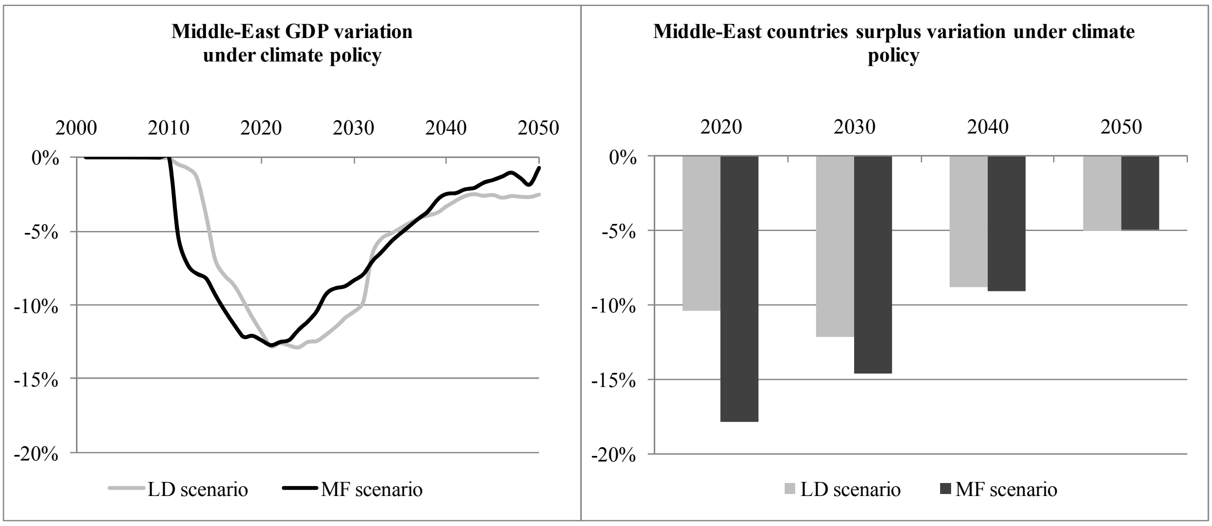
Over the short-term (2010-2020), the contraction of economic activity is very significant in the two scenarios, reaching 12% GDP losses. This is due, as in the other countries, to the succession of sharply upward-oriented carbon prices which are necessary to foster low carbon trajectories in due time under imperfect foresight. The economic effects of this carbon price are magnified by inertias on the renewal of technologies and end-use equipments which inhibits the capacity of producers and consumers to escape the increase of their energy bill.

The carbon-intensity of production thus remains high, leading to an important increase of production costs and final prices, undermining competitiveness of non-energy sectors and consumers' purchase power. Those effects combine to generate a drop in final demand, a contraction of production, a higher unemployment and an additional weakening of households' purchase power. Although valid for any region complying to the climate coalition, these mechanisms are particularly important in Middle-East countries because of the high fossil-intensity of their production process driving high energy costs (especially when compared with labour costs), and of the high share of energy expenditures in households' budget.

Over the long run, the important drop of oil revenues in the long-run happens in parallel with a recovery phase of GDP with respect to its baseline level and macroeconomic costs in 2050 are moderate in the two scenarios (1% and 3%, respectively). This result can be analyzed as the end of the 'natural resource curse' (Sachs and Warner, 2001): high resources rents do not guarantee a sustained growth pattern if limits in the absorption capacity of the economy weakens the efficiency of re-investments in non-rent production sectors. Thus, a drop of oil revenues as experienced under a climate policy is not necessarily univocally penalizing for economic activity. For a given exogenous assumption about the current account balance[12], lower oil exportation revenues imply a lower exchange rate of local currencies and foster local production at the expense of industrial goods importation. The climate policy then induces a faster industrialization in Middle-East countries, which better prepares their economies to the post oil era. As a result, long-run economic activity is less sensitive to oil revenues and those revenues fall into a more mature production structure apt to absorb them efficiently and to ensure a more sustained economic activity. This is critically illustrated by the MF scenario, in which the same GDP levels are reached in 2050 despite 25% lower oil revenues at that time horizon.

[12] As explained in Section II, the capital balance is exogenous and exponentially reduced to zero for all regions at the end of the simulation period.

Figure 2.4. *Variations under climate policy of (a) GDP [left-hand panel] and (b) surplus [right-hand panel]*



To measure the socio-economic consequences of these economic losses, we go beyond the analysis of GDP, which, although a good measure of economic activity, is only a partial indicator of welfare. We consider in Figure 2.4b the variation of households' surplus, which is commonly admitted as a better indicator in the context of important changes of the structure of prices (Bernard and Vielle, 2003) and which isolates the consequences of the climate policy on households.

The general surplus trend is similar to GDP variations with important losses in the short-term and a long-term partial recovery. But, contrary to GDP trends, the magnitude of the effects is notably different according to the pricing scenario. This is particularly true in the short-term where welfare losses are amplified under MF scenario (up to 18%). In this case indeed, the introduction of a carbon price has an important relative effect on end-use prices compared to low baseline oil price and hence undermines households' purchase power. This effect also happens under a LD scenario but with a lower relative magnitude since high short-term oil prices already affect households' purchase power in the baseline, making the relative losses more moderate (around 10%).

At a longer term horizon, the recovery of surplus levels proves to be less important than when considering GDP and welfare is still 5% lower than in the baseline in 2050. Indeed, it is more difficult to decrease the oil-intensity of final demand than for global activity because of inertias limiting the changes of consumption patterns. The more important of these inertias in terms of carbon emissions (and hence of costs of climate policy) is associated to constrained mobility needs for daily travels (essentially, commuting and shopping), which force the consumption of fossil fuels[13] without increasing welfare levels. Since these mobility needs are dependent on many other determinants than energy prices (e.g., prices of land and real estates, infrastructure availability, housing preferences), they are not notably changed by the introduction of the carbon price, even in the long term, and create an important burden on households' budget in the final phase of carbon price increase, hence limiting the positive effect of structural change away from oil.

IV. Climate policy and Monetary Compensations (MC) for oil producers

1- Which monetary compensation for oil producers?

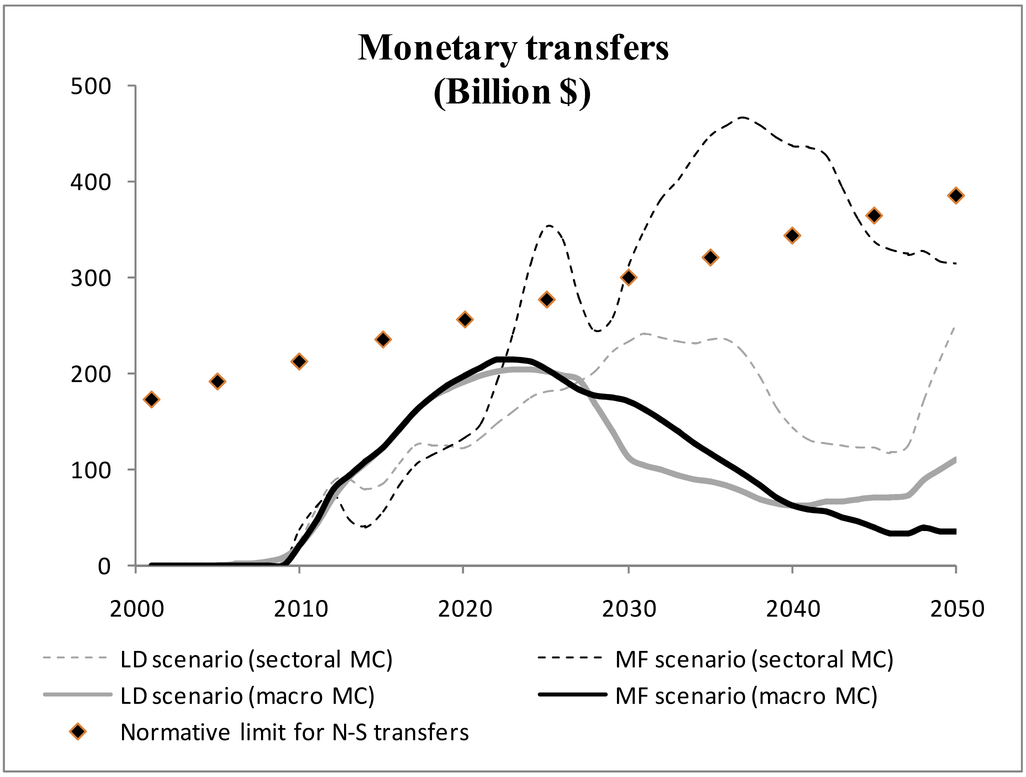
The previous section has demonstrated that, even if the costs of the climate policy are reduced when considering economic activity instead of oil revenues, they remain high enough (especially in the transitory period) to cast doubt on the participation of Middle-East countries to a climate coalition. It rather suggests that this participation may be conditional upon the introduction of monetary transfers (mostly from OECD countries) to compensate the adverse impacts of the climate policy. The amount and time-profile of those transfers (Figure 2.5) could aim at compensating the losses of oil exportation revenues (sectoral MC) or the drop of economic activity (macroeconomic MC).

The first option compensates the drop of exportation revenues due to the climate policy (dotted lines in Figure 2.5). Although this approach may be considered as the more natural way to assess the adverse impacts of climate policies on oil exporters, two elements casts doubt on the actual implementation of such transfers. The first simple one is related to their political acceptability, as measured by the burden it imposes on OECD. To appreciate this dimension, it is sufficient to observe that the transfers would be close or even exceed the normative 0.7% of GDP that is set as the UN target for total North-South transfers for development (diamond points in Figure 2.5). The second issue relates to the sensibly different time-profile of the temporal distribution of monetary compensations and of macroeconomic losses due to the climate policy. Indeed, monetary compensations remain relatively moderate until 2020 at the period when Middle-East countries experience the more intense

[13] in absence of large scale fossil-free options for transport

macroeconomic losses, whereas they are very important in the long-run despite the catch-up of baseline levels at that time horizon. This mismatch is likely to affect the efficiency of sectoral monetary transfers to limit the costs of the climate policy for oil exporters, especially during the short-term period.

Figure 2.5. *Monetary compensations for Middle-East countries (Billion $)*

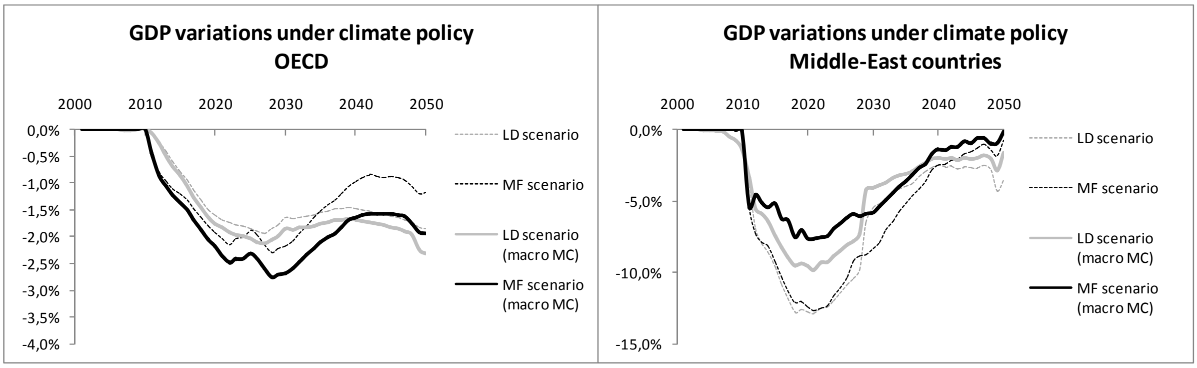


For those two reasons, we will not consider more in-depth sectoral monetary transfers but consider instead that the transfers should be assessed to compensate economic activity losses, as measured by the reduction of Middle-East's GDP due to the climate policy (solid lines in Figure 2.5). They still represent a relatively important amount (around 0.24% of OECD GDP on average over 2010-2050 in both scenarios), but they always remain below the 0.7% benchmark and leave room for other North-South transfers. Perhaps more importantly, their time-profile is very different from the transfers based on oil exportation revenues: they are important in the short-term but become very modest over the long-term.

2- General equilibrium effects of monetary transfers

Macroeconomic monetary transfers then seem relevant to address the two limitations identified above in the case of sectoral transfers, as confirmed by their effect on macroeconomic costs in OECD and Middle-East countries (Figure 6).

Figure 2.6. *Relative GDP variations under climate policy (a) in OECD [left-hand panel]; (b) in Middle-East countries [right-hand panel]*



On the one hand, although the transfers unsurprisingly increase the costs of the climate policy in OECD, they do not induce excessive burdens in those oil importing countries. In the long-term losses never exceed 2.5% of GDP (Figure 6a). On the other hand, the time-profile of monetary transfers has the advantage of supporting economic activity in Middle-East countries during the short-term period 2010-2030, when the most important losses are experienced. They will help reducing the transitory costs suffered in Middle-East countries down to 7.5% of GDP (in MF scenario) or 9.5% (in LD scenario) instead of 12.5% (Figure 6b).

Some might have expected the costs reductions permitted by monetary transfers to be higher and to cover the major part of climate policy-induced losses. However, the above results demonstrate that compensations mechanisms are not straightforward when accounting for general equilibrium effects. To investigate more in-depth the mechanisms at play, we decompose the economic consequences of monetary compensations on Middle-East countries according to the three components of GDP, namely consumption, investments and trade balance (exportations minus importations). For each of them, we consider an “efficiency index” defined by the ratio of variations due to monetary compensation over the amount of transfers: a value *x* means that 1$ in monetary compensation brings *x*$ of economic activity.

Table 2.2. *"Efficiency index" of monetary transfers*

	Total	Consumption	Investments	Trade	*of which:*	
					Oil	*Industry*
LD scenario	0.21	0.76	0.15	-0.70	-0.02	-0.56
MF scenario	0.63	0.62	0.06	-0.05	0.11	-0.17

Table 2.2 shows that the total net effect of monetary compensation remains positive, as captured by the positive value of the total efficiency index in both scenarios (0.21 and 0.63 in the LD and MF scenario, respectively). However, these values remain significantly lower than one, meaning that the net effect of monetary transfers on economic activity is limited by general equilibrium interactions. This result explains why Middle-East countries still experience relatively important macroeconomic losses despite the monetary transfers, especially in the LD scenario (Figure 6b). Let us now consider the three components of economic activity.

Monetary transfers foster local activity as captured by the positive (and close to one) sum of consumption and investment index (0.91 in the LD scenario and 0.68 in the MF scenario). These results are confirmed by lower households' surplus losses (see Figure 8a in Section V). Indeed, the additional capital available in Middle-East helps better preparing their economy to the post-oil era through additional early investments in industrial production capacities, and then contributes to support long-term consumption. This effect is particularly important in the LD scenario since the additional capital helps to offset the redirection of investments from industrial to oil sectors fostered by high early oil price.

The positive effect on local economic activity happens at the expense of trade balance as captured by the negative value of the trade index in both scenarios (-0.70 in the LD scenario and -0.05 in the MF scenario). We distinguish the effects on the two main traded sectors, oil and industry.

First, monetary compensations affect Middle-East's oil exportations through two opposite effects on OECD's oil demand: (i) lower OECD households' disposable income, thus decreasing oil demand; (ii) lower capital availability for OECD investments, slowing down cumulative technical change towards oil-free pathways and ensuring higher long-term oil demand. This latter effect is particularly important under the MF scenario, where low oil prices concur in giving poor incentive for oil-free technical change. This explains why the effect of monetary transfers on oil exportations is globally positive in this scenario, as captured by the positive value of the oil exportation index (+0.11). On the contrary, the former effect slightly dominates in the LD scenario, leading to a small negative effect of monetary transfers on oil exportations (-0.02).

Also, monetary transfers affect the competitiveness of Middle-East's industry on international markets. Indeed, under exogenous assumption about the current account balance[14], the inflows of external capital due to monetary compensations force a degradation of the trade balance and a loss of competitiveness for Middle-East's industrial exportations. This negative effect is particularly important in the LD scenario, where high short-run oil prices delay investments in industrial production capacity hence further undermining its competitiveness over the long term.

V- What if Middle-East countries do not participate to the climate policy?

The previous section has demonstrated that the introduction of monetary compensations raises concerns about both their efficiency in reducing the costs of the climate policy in Middle-East countries and the political acceptability of associated transfers in OECD countries. An alternative would be to consider that, in absence of monetary compensations, Middle-East countries exit the climate coalition with the rationale of avoiding the costs associated to a local carbon constraint.

1- Is the withdrawal from the climate coalition efficient? Macroeconomic effects in Middle-East countries

This withdrawal would force other countries to accept an additional burden of carbon abatements and to adopt higher carbon prices than those experienced when monetary compensations ensure the participation of Middle-East countries (Figure 2.7).

[14] See Section II.

Figure 2.7. *Carbon price under the no participation case ($/tCO$_2$)*

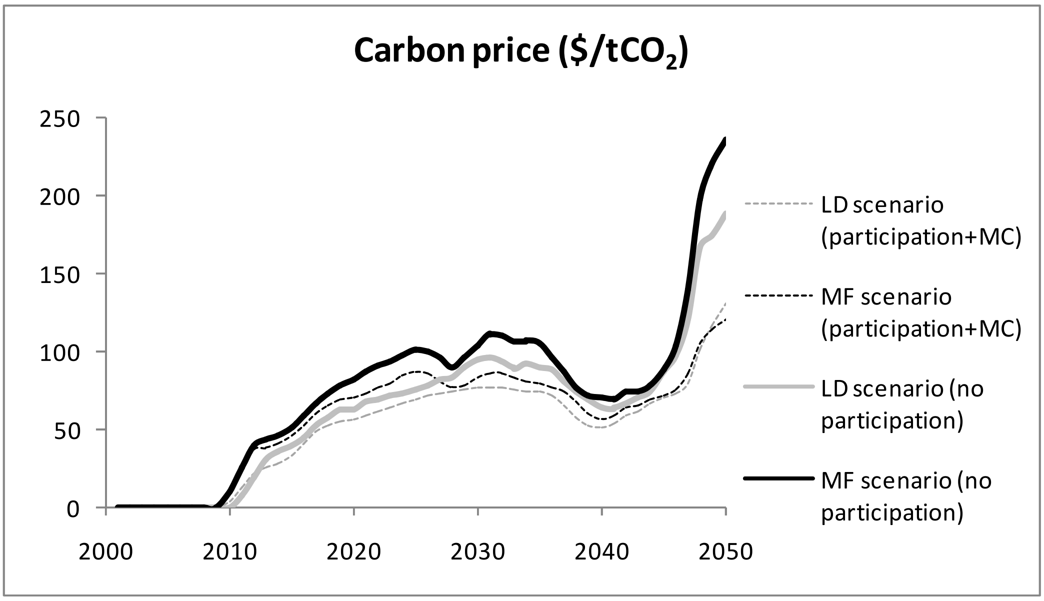
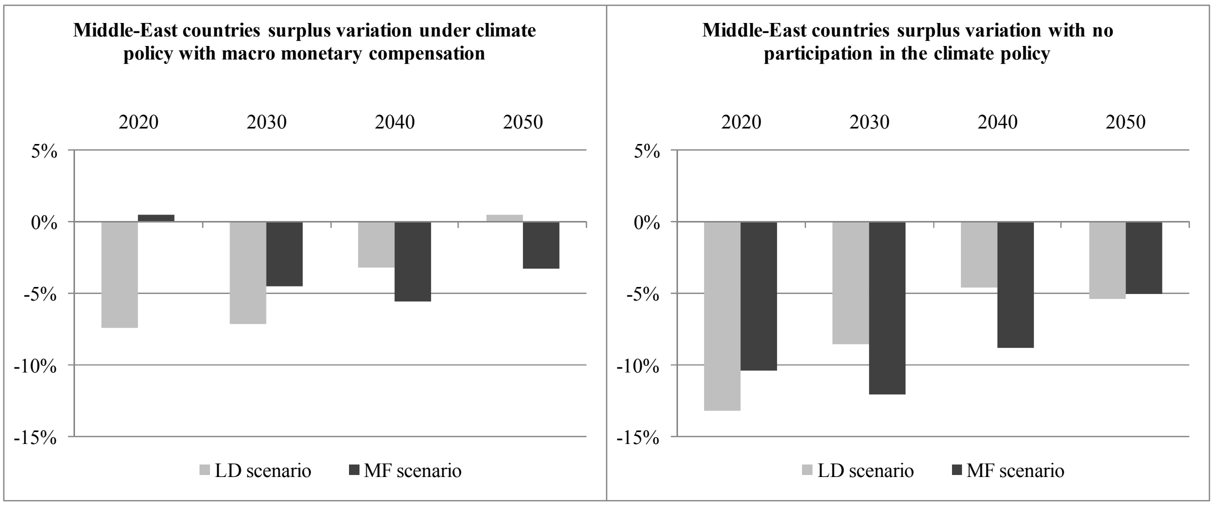


Figure 2.8 depicts Middle-East surplus losses between the baseline and (a) a global climate agreement in which they receive monetary compensations for their GDP losses and (b) a global climate agreement from which they withdraw. It shows that in both cases, the benefits in terms of purchase power and industrial competitiveness reduce the surplus losses due to the climate policy (see Figure 4b for a comparison). Residual losses under the no participation case are essentially due to the reduction of oil exportation revenues because additional carbon reductions in oil-importing countries come with a decrease of oil demand.

These losses follow very different time profiles according to the oil pricing scenario. Under a MF scenario (i.e. with low short-term oil price), the withdrawal from the climate policy helps softening the important short-term costs by avoiding the negative effects of carbon pricing on the local very oil-intensive domestic economy (surplus losses are reduced from 18% to 10% in 2020). This effect is not present in the LD scenario, since patterns of consumption are less oil-dependent due to high oil prices. In the middle to long-term, on the contrary, exiting the climate policy is more beneficial for Middle-East countries under the LD scenario (losses are reduced from 13% to 7% in 2030). Indeed, the sensibly higher carbon price (up to 100$/tCO2 around 2030) favors oil against coal in the rest of the world, which contributes to compensate the decrease in oil demand triggered by a long period of high oil prices.

Figure 2.8. *Surplus variations in Middle-East between baseline and climate policy scenarios, under (a) participation and monetary compensation [left-hand panel]; and (b) no participation [right-hand panel].*



Furthermore, Figure 8 shows that from a welfare point of view Middle-East countries would prefer to receive monetary compensations than to leave the climate coalition, whatever the oil pricing strategy. Indeed, monetary transfers significantly reduce the effect of the carbon tax on households' consumption, while in the no participation case the reduction of oil exportations revenues strongly impacts local economies.

2- Are the monetary compensations acceptable? Macroeconomic effects in OECD countries

The next question is whether or not OECD countries might accept to give monetary compensations to Middle-East countries, rather than seeing them out of the climate coalition.

In Section IV, Figure 6a showed that, from a macroeconomic point of view, monetary compensations might be acceptable by OECD countries. In a LD scenario, however, Figure 6 shows that monetary compensations increase between 2040 and 2050, together with a steep rise of the carbon price. It is thus interesting to assess the acceptability of these transfers compared to a withdrawal of ME countries in terms of households' welfare. To do so, we consider the additional OECD surplus variations due to the monetary compensations or the withdrawal of Middle-East countries) with respect to the benchmark case of section III (Table 2.3). A significantly negative value means that losses are important compared to the case of an unconditional participation of Middle-East countries.

This is the case in the short- and medium-term as measured by around 1% losses of OECD households' surplus; but these values are somewhat similar under all oil pricing scenario or the participation of Middle-East countries to the climate coalition. This means that the short to medium term welfare costs are not significantly different whether they are due to monetary compensations or to a withdrawal of ME countries. By considering this outcome, OECD countries should be indifferent between agreeing to give monetary compensations to Middle-East producers and taking the risk of seeing them out of the climate coalition.

Table 2.3. *"Acceptability indicator" of monetary transfers vs risk of ME withdrawal (OECD additional surplus variation compared to the benchmark, i.e. the case of an unconditional participation of Middle-East countries). The last two lines indicate whether monetary compensations create losses (-) or gains (+) for the OECD compared to a withdrawal of ME countries. (0) is for differences lower than 0.1% and (--) for differences higher than 1%.*

		2010-2020	**2020-2030**	**2030-2040**	**2040-2050**
Monetary compensations	LD scenario	-1.3%	-1.4%	-0.5%	-1.6%
	MF scenario	-1.1%	-1.4%	-0.9%	-1.8%
No participation of ME countries	LD scenario	-0.9%	-1.0%	-0.6%	+0.3%
	MF scenario	-1.0%	-1.5%	-1.5%	-2.2%
Comparison	**LD scenario**	**-**	**-**	**0**	**--**
Mon. comp. *vs* no particip	**MF scenario**	**0**	**0**	**+**	**+**

However, the participation of Middle-East countries is critical for OECD economies at a longer term horizon (2050), with opposite effects according to the pricing scenario considered. If short-term oil prices are high (LD scenario), the withdrawal of Middle-East from the climate coalition has a long-term benefitial effect on OECD households' surplus (+0.3%): it forces higher short-term carbon prices which act as an early incentive to accelerate carbon-free technical change with the twofold effect of limiting the surge of long-term carbon prices and making the economy less vulnerable to energy shocks. On the contrary, the option with monetary transfers is very costly (-1.6%): the constraints they impose on the availability of investments and the lower carbon prices contribute to slower oil and carbon-free technical change in OECD and hence make households more vulnerable to energy variations (including carbon price). These effects are also present with low short-term prices (MF scenario), but with different magnitudes. In particular, the increase of carbon prices is particularly steep if Middle-East does not participate (up to 250$/tCO2), which is too sudden to be absorbed by the economy and causes important surplus losses (-2.2%)

This demonstrates a close interplay between geopolitical issues on oil pricing and climate negotiations. Indeed, if oil prices are maintained at a high level (LD scenario), OECD would not be incline to concede monetary compensations in exchange for the compliance of Middle-East producers, and they would instead prefer Middle-East countries to exit the coalition since this notably lowers their long-term losses. On the contrary, a short-term fall of oil prices would lower Middle-East short-term profit but would put them in a better position to negotiate monetary compensations. Indeed, in that case the OECD would risk high long-term losses if Middle-East withdrew from the coalition. Furthermore, under a MF scenario short-term welfare losses would be completely offset by monetary compensations in Middle-East countries.

VI. Conclusion

This paper analyses the interactions between oil markets and the macroeconomy when a climate policy is implemented. Three main messages can be derived from the modeling exercise.

First, the aggregate costs incurred in Middle-East countries under a global agreement on ambitious climate policy may remain limited, but the time profiles demonstrates the risk of one or two transitory decades with important slowing down of economic activity and households welfare. This analysis casts doubt upon the participation of Middle-East countries in absence of compensations designed to make this transition acceptable.

This paper envisages two types of such architecture based on the compensation of either oil exportation revenues or macroeconomic losses by OECD countries. In the latter case, transitory losses are reduced by 1/3 without inducing an excessive burden in OECD countries, but losses are not cancelled by the compensations.

Finally, the cost of Middle-East exiting the climate coalition in absence of agreement on the monetary transfers remains moderate and essentially depends on oil pricing trajectories in the long-term. At this time horizon, the results demonstrate a close dependence between geopolitical dimensions of oil markets and climate negotiations. Indeed, high oil prices put Middle-East countries in an unfavorable position to obtain monetary compensations in climate negotiations, whereas low short-term oil price trigger such important long-term losses in OECD that these countries would be more encline to accept monetary compensations to gain the compliance of Middle-Ease countries to the climate coalition.

References

Armington PS (1969). A Theory of Demand for Products Distinguished by Place of Production. *IMF, International Monetary Fund Staff Papers* 16: 170-201.

Barker T, Bashmakov I, Bernstein L, Bogner JE, Bosch PR, Dave R, Davidson OR, Fisher BS, Gupta S, Halsnæs K, Heij GJ, Kahn Ribeiro S, Kobayashi S, Levine MD, Martino DL, Masera O, Metz B, Meyer LA, Nabuurs G-J, Najam A, Nakicenovic N, Rogner HH, Roy J, Sathaye J, Schock R, Shukla P, Sims REH, Smith P, Tirpak DA, Urge-Vorsatz D, Zhou D (2007): Technical Summary. In: *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, L. A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Barnett J, Dessai S, Webber M (2004). Will OPEC lose from the Kyoto Protocol? Energy Policy 32(18), 2077--2088.

Barnett J, Dessai S (2002). Articles 4.8 and 4.9 of the UNFCCC: adverse effects and the impacts of response measures. Climate Policy 2(2-3), 231--239.

Bartsch U and Muller B (2000). Impacts of the Kyoto Protocol on fossil fuels. *In*: Bernstein, L., Pan, J. (Eds.), Sectoral and Economic Costs and Benefits of GHG Mitigation. Intergovernmental Panel on Climate Change, Bilthoven.

Bernard A and Vielle M (2003). Measuring the welfare cost of climate change policies: A comparative assessment based on the computable general equilibrium model GEMINI-E3. *Environmental Modeling and Assessment* 8: 199–217.

Berg E, Kverndokk S and Rosendahl KE (1997a). Gains from cartelisation in the oil market. Energy Policy 25 (13), 1075--1091.

Berg E, Kverndokk S and Rosendahl KE (1997b). Market Power, International CO_2 Taxation and Petroleum Wealth . The Energy Journal 18(4), 33--71.

Bresnahan TF, Ramey VA (1993). Segment Shifts and CapacityUtilization in the U.S. Automobile Industry. American Economic Review Papers and Proceedings 83(2): 213-218.

Davis S, Haltiwanger J (2001). Sectoral Job Creation and Destruction Responses to Oil Price Changes. Journal of Monetary Economics 48: 465-512.

Finn MG (2000). Perfect competition and the effects of energy price increases on economic activity. Journal of Money, Credit and Banking 32(3): 400-416.

Ghanem S, Lounnas R, Brennand G (1999). The impact of emissions trading on OPEC. OPEC Review 23(2), 79--112.

Hamilton JD (2008). Oil and the Macroeconomy. The New Palgrave Dictionary of Economics Intergovernmental Panel on Climate Change (IPCC), 2001. Climate Change (2001): The Scientific Basis. Cambridge University Press, Cambridge.

Holland S (2008). Modeling Peak Oil. The Energy Journal 29(2): 61-80.

Johansson D, Azar C, Lindgren K and Persson T (2009). OPEC Strategies and Oil Rent in a Climate Conscious World. The Energy Journal 30 (3): 23-50.

McKibbin W, Ross T, Shackleton R, Wilcoxen P (1999). Emissions trading, capital flows and the Kyoto Protocol. *In*: Weyant, J.(Ed), The Energy Journal, special issue.

Persson T, Azar C, Johansson D, Lindgren K (2007). Major oil exporters may profit rather than lose, in a carbon-constrained world. Energy Policy 35(12), 6346--6353.

Rotemberg J, Woodford M (1996). Imperfect Competition and the Effects of Energy Price Increases. Journal of Money, Credit, and Banking 28: 549-577.

Schlessinger JR (1989). Energy and geopolitics in the 21st century. World Energy Conference 14th Congress Montreal, Quebec, September 20th, 1989.

van Vuuren DP, den Elzen MGJ, Berk MM, Lucas P, Eickhout B, Eerens H, Oostenrijk R, (2003). Regional costs and benefits of alternative post-Kyoto climate regimes: comparison of variants of the multi-stage and per capita convergence regime. Report 728001025, RIVM, Bilthoven, the Netherlands.

WBGU (2003). Climate Protection Strategies for the 21st Century: Kyoto and Beyond. German Advisory Council on Global Change, Berlin, Germany.

.

Chapter 3

Infrastructures, Technical Inertia and the Costs of Low Carbon Futures under Imperfect Foresight

Chapter 3* examines the global and long-term dimensions of the interplay between carbon prices, oil prices and the macroeconomy by analyzing the transition costs of moving towards a low carbon society. We emphasize the consequences on mitigation costs of considering the interplay between a) technical systems inertia, including slow infrastructure turnover in transportation and construction; and b) imperfect foresight influencing investment decisions. To this end, the hybrid general equilibrium modeling framework IMACLIM-R is employed as it allows for transitory partial adjustments of the economy and captures their impact on the dynamics of economic growth. The modeling exercise quantitatively emphasizes the i) specific risks that the interplay between inertia and imperfect foresight leads to high macroeconomic costs of carbon abatement measures; ii) opportunities of co-benefits from climate policies permitted by the correction of sub-optimalities in the reference scenarios. This chapter draws insights for the framing of future climate architectures by studying the role of measures that act complementarily to carbon pricing in the transport sector. In particular, reallocating public investment towards low-carbon transport infrastructure significantly reduces the overall macroeconomic costs of a given GHG stabilization target and even creates the room for long-term net economic benefits from climate policies.

---

* This chapter is the reproduction of
Waisman H, Guivarch C, Grazi F and Hourcade JC (2011). The Imaclim-R Model:Infrastructures, Technical Inertia and the Costs of Low Carbon Futures under Imperfect Foresight, *accepted in Climatic Change*

Economic analysis of climate policy faces a paradox. The literature suggests that the macroeconomic cost of achieving stringent GHGs concentration targets would be moderate[1], but most countries remain reluctant to adopt whereas ambitious climate policy. This may be because the few percentage points of GDP losses, translated in billion dollars, represent a prohibitive cost for decision makers; but part of the paradox may also lie in an often disregarded caveat of the IPCC report:

> "Most models use a global least cost approach to mitigation portfolios and with universal emissions trading, assuming transparent markets, no transaction cost, and thus perfect implementation of mitigation measures throughout the 21st century" (IPCC, 2007, Box SPM.3).

This caveat actually points out the deficit of information about the transition to a low-carbon future in a second-best world. One can argue that imperfect foresight, incomplete markets and institutional failures will lead to higher costs that those reported so far, or, conversely, that non optimal baselines offer opportunities for relative gains under climate policy. This is the major finding of Barker and Scrieciu (2010) with the macroeconometric E3MG model.

The IMACLIM-R model (Sassi et al., 2010) contributes to this debate by incorporating some features of second-best economies in a Computable General Equilibrium model, able to represent important structural and technical change over a century. Its main specificity is to endogenize transitory adjustments of an economy constrained by the interplay between choices under imperfect foresight and the inertia of technical systems. Imperfect foresight is a consequence of *a*) uncertainty about future relative prices, final demand and investments profitability, *b*) "noises" coming from signals other than energy prices (informal economy, prices of land and real estate) and *c)* non-economic determinants of public decisions in transportation and urban planning. Of little importance in a flexible world, imperfect foresight becomes crucial when non-optimal choices cannot be corrected frictionless because of inertias on capital stocks and behavioral routines.

[1]The IPCC (2007) reports costs between small GDP gains and lower-than-5.5% losses of global GDP in 2050 for stabilization targets between 445 and 535 ppm CO_2-eq (Table SPM.6). The ADAM project (Edenhofer et al., 2010b), extends the estimates to 2100 and finds aggregate costs below 2.5% of global GDP for 400 ppm CO_2-eq targets.

The objective of this paper is to show how this interplay between imperfect foresight and inertia explains the peculiar shape of the mitigation cost profiles found by IMACLIM-R compared to those in most models, including WITCH and REMIND in the RECIPE project (Edenhofer et al, 2010a; Luderer et al, 2010a). Section 2 sums up the overall rationale of the IMACLIM-R model and insists on the specific role of infrastructures, a typical case of rigid capital stock. Section 3 shows out the major economic reasons for high short term climate policy costs, specifically in emerging economies, and for potential long run benefits. Section 4 investigates how cost profiles change with a richer climate policy package including complementary measures to carbon pricing, specifically infrastructure policies that affect the transport sector.

I. Rationale of the IMACLIM-R modeling Structure

The IMACLIM-R model is a recursive, dynamic, multi-region and multi-sector[2] Computable General Equilibrium (CGE) model of the world economy. It allows for describing growth patterns in second best worlds (market imperfections, partial uses of production factors and imperfect expectations) through a hybrid and recursive dynamic architecture. A detailed algebraic description of this model is given in Annex A and we herein outline only those of its features that matter for the purpose of this paper.

1. A growth engine with gaps between natural and effective growth

IMACLIM-R incorporates exogenous assumptions of regional labour productivity growth and active population growth (see Annex A), which determine the exogenous 'natural' growth rate.[3] Effective growth rates are endogenously driven by labour allocation across regions and sectors at each point in time, given relative productivities and short-term rigidities (capital stock inertia, frictions in reallocating labour and wage rigidity). Aggregate capital accumulation is controlled by exogenous saving rates like in Solow (1956), but IMACLIM-R represents investment decisions under imperfect foresight. At a given date, agents have limited information about the future and shape their expectations on the basis of past and current trends (adaptive expectations). Under such semi-myopic foresight, installed capital

[2] see Annex A for regional and sectoral disaggregations.
[3] A large strand of literature has emerged after Solow (1956) that traditionally represents growth trajectories on the basis of this "natural" growth rate, which boils down to representing the global economy as characterized by a unique composite production sector operating at full employment.

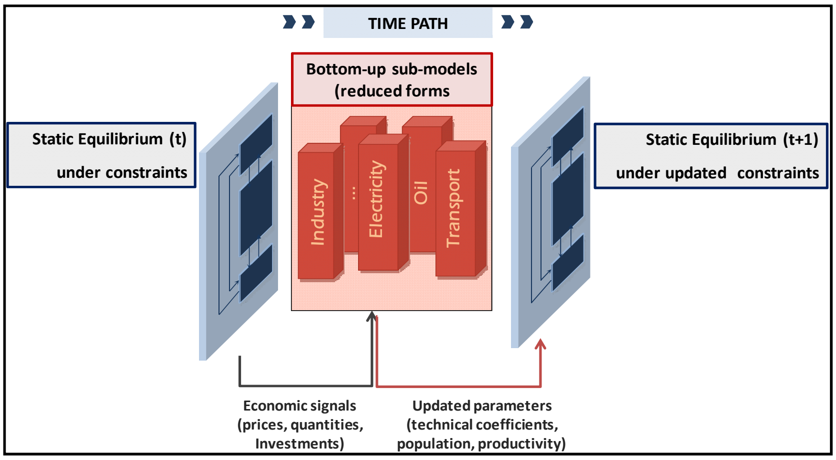
resulting from past investment decisions may not be adapted to future economic settings. However, it cannot be renewed overnight due to inertias and acts as a constraint on the adaptability to variations of economic conditions (activity levels and prices).

2. A recursive and modular architecture to endogenize technical change

The IMACLIM-R model endogenizes the rate and direction of technical change through describing the impact of investment decisions on the deployment of technical systems. The consistency of the top-down/bottom-up conversation is guaranteed by a hybrid structure representing the economy in money values and physical quantities (Hourcade et al, 2006). This dual accounting follows the Arrow-Debreu axiomatic. It ensures that the projected economy is supported by a realistic technical background (in the engineering sense) and, conversely, that projected technical systems correspond to realistic economic flows and consistent sets of relative prices. In climate policy analysis, this approach has for long been claimed as crucial for the energy goods to represent explicitly their carbon-to-energy ratio (Malcolm and Truong, 1999; Sands et *al.*, 2005). IMACLIM-R extends it to transportation as another key sector of climate analysis.

A recursive structure then organizes a systematic exchange of information between a top-down annual static equilibrium providing a snapshot of the economy at each yearly time step, and bottom-up dynamic modules informing on the evolution of technical parameters between two equilibria (Figure 3.1).

Figure 3.1: *The recursive and modular structure of the IMACLIM-R model*



The annual static equilibrium determines relative prices, wages, labour, value, physical flows, capacity utilization, profit rates and savings at date t as a result of short term equilibrium conditions between demand and supply on all markets, including energy. Utility-maximizing households base their consumption choices on both income and time constraints; the former is the sum of wages, capital returns and transfers whereas the latter controls the total time spent in transportation. Firms adapt their short term production considering fixed input-output coefficients (the average of techniques embodied in their capital stock) and decreasing static returns when capacity approaches saturation[4]. They determine their prices with a margin rate over production costs (mark-up) to capture the effect of imperfect competition.[5]

Total demand for each good (the sum of households' consumption, public and private investments and intermediate uses) is satisfied by a mix of domestic production and imports.[6] All intermediate and final goods are internationally tradable. Domestic as well as international markets for all goods are cleared (i.e. no stock is allowed) by a unique set of relative prices and this determines the utilization rate of production capacities.[7] The equilibrium values of all variables are sent to the dynamic modules to serve as a signal for agents' decisions affecting technical coefficients at $t+1$.

The dynamics of the economy is governed by endogenous descriptions of capital accumulation and technical change, given the exogenous 'natural' growth assumptions. At each year, regional capital accumulation is given by firms' investment, households' savings, and international capital flows[8]. On that basis, the across-sector distribution of investments is governed by expectations on sector profitability and technical conditions as described in

[4] Following (Corrado and Mattey, 1997), decreasing returns reflect the higher labor costs associated to extra-hour operations, costly night work and increasing maintenance works when capacity approaches saturation.

[5] The mark-ups are exogenous except in energy sector where they are endogenous to reflect (a) the market power of fossil fuel producers (b) specific pricing principles in the power sector (e.g., mean cost pricing), and (c) the different margins over the three inputs for liquid fuels production (oil, biomass, coal).

[6] For non-energy goods, we adopt Armington specifications (Armington, 1969) to capture the partial substitutability between domestic and foreign goods, while physical accounting for energy goods (in MToe) makes them fully substitutable.

[7] The partial utilization rate of production capacities allows representing operational flexibility through early retirement of those capacities which, although installed, are not used for actual production because not competitive in current economic conditions.

[8] In absence of explicit interest rate, we assume a gradual correction of current imbalances, as a standard proxy for the complex determinants of international capital flows in energy forecasting exercises (Edmonds et al, 2004; Paltsev et al., 2005).

sector-specific reduced forms of technology-rich models (referred to as Nexus modules and extensively described in Annex A).

The Nexus modules represent the evolution of technical coefficients resulting from agents' microeconomic decisions on technological choices, given the limits imposed by the innovation possibility frontier (Ahmad, 1966). They embed *a*) sector-based information of economies of scale, learning-by-doing mechanisms and saturation in efficiency progress, and *b*) expert views about the asymptotes on ultimate technical potentials, the impact of incentive systems, and the role of market or institutional imperfections. The new investment choices and technical coefficients are then sent back to the static module in the form of updated production capacities and input-output coefficients to calculate the t+1 equilibrium.

This structure comes to adopt a standard putty-clay representation with fixed technical content of installed capital, which allows distinguishing between short-term rigidities and long-term flexibilities (Johansen, 1959).

3- A specific treatment of the transport sector

The potentials of the IMACLIM-R structure have been exploited to explicit the specifics of the transport sector and its impact on energy demand. This sector, vital for economic and human development, is characterized by a strong path dependency of options, by the influence of non-energy determinants in the collective and individual behaviors (for example the spatial setting *via* location choices of both firms and households) and by the dependence upon long-lived infrastructure investments.

Transport demand is indeed affected by *i*) the attractiveness of alternative modes and influencing the modal choice of individuals, *ii*) households' mobility needs and their average (travel and commuting) distances, and *iii*) the spatial organization of production and its associated freight transport needs. IMACLIM-R incorporates these features in three ways:

i. The relation between transportation infrastructure, mobility demand and modal choices is captured in the maximization of households' utility where saturation of infrastructures cause speed decreases when normal load conditions are exceeded. Then, investments in transport infrastructures determine the efficiency of the different transport modes and, hence, the allocation of travel time budget across modes of different efficiencies.

ii. The utility demand for mobility includes households' constrained mobility (for commuting and shopping). It does so through a "basic need' level which depends essentially on location and infrastructure constraints (residential areas, work centers, transport infrastructures). Non directly sensitive to fuel prices variations, they capture location and infrastructures choices, including urban policies aimed at limiting urban sprawl.

iii. The freight transport content of production processes is represented by explicit input-output coefficients. The absence of decoupling between production and transport (constant input-output coefficient) corresponds to pursuing current trends of transport-intensive production; in alternative scenarios we also consider a progressive decrease of the freight content of production in a way to represent changes in producers' choices on the supply chains, relocation of production infrastructures (more vertically integrated, and spatially closer to markets) and a moderation of "just-in-time" processes.

II. Time Profiles of Climate Policy Costs

We herein analyze the specifics of cost profiles of a global climate policy in IMACLIM-R. In section 1, we test their robustness to parameter uncertainty. In section 2, we use a simple analytical demonstration to clarify the short-term economic determinants of costs and their evolution over time. In section 3, we demonstrate why the long distance race between technical change and inertia is the major cause of important short-term losses and allows for long-term catch-up of baseline levels (and even some benefits).

To conduct these analyses, our numerical experiments will encompass (see Appendix)

A. three assumptions on *Oil and Gas supply :* (A-1) assumes moderate limitations on medium-term oil supply in line with conservative estimates on the amount and distribution of oil reserves, whereas (A-2) considers lower oil reserves. We add a third variant, an even more pessimistic case, where not only are resources low, but also geological constraints on capacity deployment forces to an accelerated decline of global production (A-3).

B. two assumptions on the *Substitutes to oil*: *(*B-2) corresponds to a faster and deeper market diffusion of substitutes to oil (biofuels and Coal-To-Liquid) than (B-1).

C. two assumptions on *Demand side technological change* : (C-2) refers to a faster diffusion of decarbonization and energy efficiency technologies than in (C-1).

Using the same assumptions on regional natural growth, we obtain very similar mean GDP growth under all combinations of parameter assumptions defining the $3 \times 2^2 = 12$ 'BAU' scenarios for the 2010-2100 period[9]. These scenarios feature a wide dispersion of CO_2 emissions from 30 to 69 $GtCO_2$ in 2100 (see Table 1), which lie in the middle of the 15-135 $GtCO_2$ range of the SRES and post-SRES scenarios, in 2100 (Barker et al., 2007, Figure TS7).

Table 3.1: *Mean annual real GDP growth in BAU scenarios, for the world and a selection of regions (average values in bold, full range of variations across scenarios into brackets).*

	World	USA	Europe	China	India
mean annual growth (2010-2050)	**2.1** [2.0;2.2]	**1.7** [1.6;1.8]	**1.6** [1.4;1.7]	**4.0** [3.8;4.1]	**4.3** [3.8;4.5]
mean annual growth (2010-2100)	**1.7** [1.6;1.7]	**1.8** [1.7;1.8]	**1.4** [1.3;1.4]	**2.4** [2.3;2.5]	**2.9** [2.7;3.0]

Since the objective is to analyze the mechanisms of cost formation, we worked under an identical CO_2 emission trajectory for all stabilization scenarios over the period. We thus set aside the question of the intertemporal flexibility for allocating emission reductions for the same carbon budget that could affect near-term mitigation costs by postponing emission reductions ("when flexibility"). The trajectory is chosen in category III of IPCC scenarios corresponding to a stabilization target of 440-485 ppm CO_2: global CO_2 emissions peak in 2017 and are decreased by 20% and 60% with respect of 2000 level in 2050 and 2100, respectively (Barker et al., 2007, Table TS2). For the purpose of this exercise, we also exclude international redistribution of tax revenues[10] and the model endogenously calculates the world carbon tax to be imposed to meet the emissions constraint at each point in time.

1- 'Carbon price-only' policy: a time profile robust to uncertainty

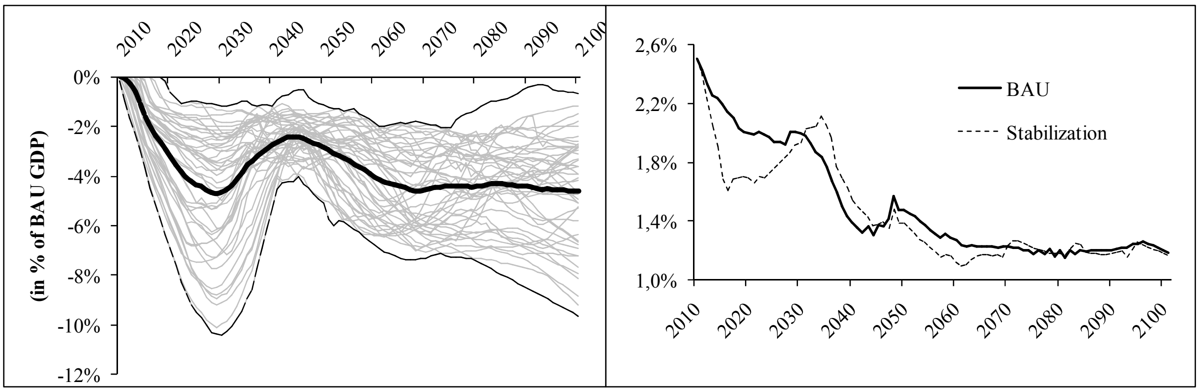
Figure 3.2(a) displays global GDP variations in stabilization scenario targets compared to the Business As Usual (BAU) situation, for all of the 12 'future worlds', the bold black line giving the average costs of these scenarios[11]. Figure 3.2(b) compares the average growth rates in BAU and stabilization scenarios.

[9] Note that the reference scenario from the RECIPE model comparison exercise (Luderer et al, 2010a) is not included within these BAU scenarios. Indeed, specific exogenous forcing of the model were introduced in the model comparison exercise in order to make the reference scenarios comparable across models. For example, an exogenous oil price trajectory was used, but is not used in the BAU scenarios from this article.

[10] The RECIPE project investigates the consequences of regional differences in carbon tax (Jakob et al, 2010) and the effects of alternative rules for quota allocation among regions (Luderer et al, 2010b)

[11] This average value attributes equal importance to each 'future world' and should not be intended as a best-guess estimate. It is displayed to identify the general trends of the variables under consideration, independently of their variability across scenarios.

Figure 3.2: *Global GDP variations between stabilization and BAU scenarios, over the 2010-2100 period* [left-hand panel]; *Average GDP growth rate across all BAU (solid line) and stabilization (dotted line) scenarios* [right-hand panel]



With a 3% discount rate, discounted mitigation costs over the period 2010-2100 range from 1% to 4.6% across scenarios, but this aggregated value masks critical issues revealed by the time profile of costs. Actually, despite differences in the magnitude of GDP variations across scenarios, four phases of GDP losses (Figure 3.2(a)) and carbon price (Figure 3.3) can be identified for all of them:

(i) substantial transitory costs during the first two decades of stabilization with lower growth rates than in reference scenarios (but never an absolute decrease of GDP in any region).[12] These costs are associated with a sharp initial increase of the carbon price.

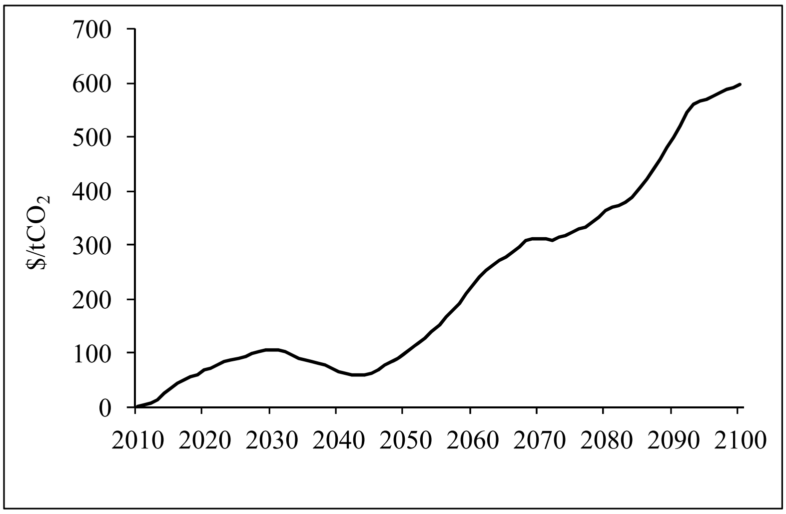
(ii) a medium-term (about fifteen years) GDP catch-up with higher growth rates under a climate policy than in the reference scenario; this phase is associated with a decline in the carbon price ending around 2045,

(iii) a second phase of significant GDP loss in the stabilization scenario from 2045 to 2070 associated with a second phase of steep carbon price increase,

(iv) a long-term regime in which, on average, the continuous increase of average carbon prices do not trigger significant GDP loss, as if the economy were adapted to a regime with ever increasing carbon prices. In fact, this average stabilization of GDP losses hides a divergence across scenarios, between optimistic scenarios with a slow economic catch-up towards baseline levels and pessimistic ones with continuing departures from these levels.

[12]IPCC reports global GDP losses between 0.2% and 2.5% in 2030 (IPCC, 2007, Table SPM.4), whereas we obtain a range between 1% and 9.5%, with the average value around 4% (see Figure 2).

Figure 3.3: *Average carbon tax ($/tCO$_2$)*



The magnitude of the above mechanisms features some significant regional difference, since the costs remain moderate in developed countries, but are extremely high in the rest of the world (Figure 3.4 and Table 3.2). This is particularly true in the short-term, in which developing countries face, on average, as high as 10% transitory losses around 2030 (against only 1% at the same time horizon in developed countries).

Figure 3.4: *Average GDP variations between stabilization and BAU scenarios in developed countries, developing countries and the world, over the period 2010-2100*

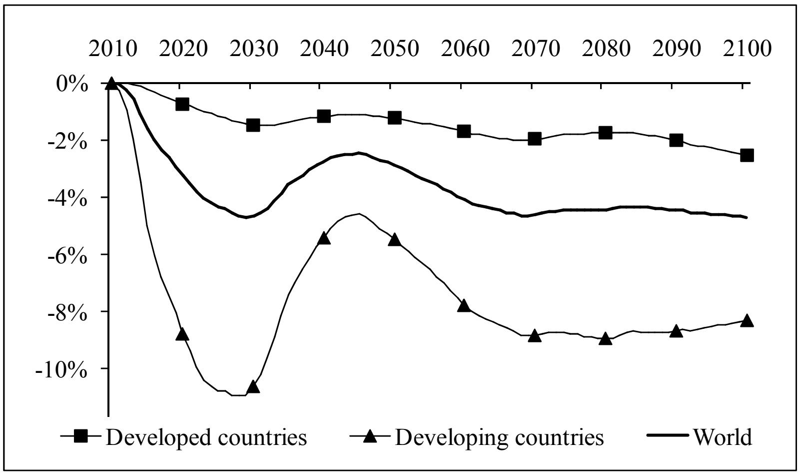


Table 3.2: *GDP losses (a negative value represents an actual gain) between stabilization and BAU scenarios for a selection of major countries and regions, in 2030, 2050 and 2100 (average values in bold, full range of variations across scenarios into brackets).*

	2030	2050	2100
USA	**1.2** [0.1; 2.4]	**1.4** [-0.1; 2.4]	**2.3** [0.2; 5.3]
Europe	**0.8** [-0.4; 2.4]	**0.7** [-0.8; 1.8]	**1.6** [-0.7; 4.5]
China	**16.8** [5.3; 33.9]	**6.9** [4.5; 12.1]	**10.5** [-3.9; 21.4]
India	**13.5** [3.6; 21.6]	**7.5** [4.0; 15.0]	**10.9** [0.8; 21.1]

2. Drivers of mitigation costs: an analytical detour

The economic drivers of these non conventional time profiles can be derived from a stylized model which incorporates the core specificities of the static equilibrium of the IMACLIM-R model. Let us assume an economy producing a composite good Q with energy and labour as input factors, a mark-up price equation to represent imperfect competition and a wage-curve to capture labour market imperfections. A simple analytic derivation (see Annex A) gives an explicit expression for mitigation costs when a tax on energy τ_E (taken as a proxy for a carbon tax) is levied, as measured by production variations ΔQ :

$$\frac{\Delta Q}{Q_0} = \frac{z_0}{1-z_0} \cdot \left[1 - \left(1 - \frac{p_E . e}{w_0 l} \tau_E \right)^{-\frac{1}{\alpha}} \right] \tag{1}$$

In equation (1), e and l are the unitary energy and labour requirements for production, p_E the price of energy, w_0 and z_0 the wage rate and unemployment level in the absence of taxation and α is the elasticity of the wage curve (the higher α, the more flexible the labor markets).

The remarkable lesson of equation (1) is that the energy parameters are not the only drivers of the costs induced by a given level of carbon taxation. Part of the magnitude of the costs is driven by the macroeconomic effects on labour markets controlled by the elasticity of the labor market α and the unemployment rate z_0 [13]. Another part is driven by the energy costs /

[13] These effects are analyzed more in-depth in (Guivarch et al.,2010)

labor costs ratio $\left(\frac{p_E.e}{w_0 l}\right)$: the countries the more adversely affected by higher energy prices are those in which the energy share in production costs is high and the salaries share is low.

The dynamic effect of the carbon tax then depends on the interplay between the pace of a) changes in labour markets driving wage adjustments, b) technical change favouring lower labour and energy inputs for production, c) energy price and carbon tax increases.

Over the short-term, the decline of the energy costs / labor costs ratio is constrained by inertias on the evolution of the ratio *e/l* and on the absolute increase of wages. This justifies high magnitude of the costs.

Over the long term, induced technical change accelerates the decline of the energy content of production and slows down the increase of energy costs; labour costs increases thanks to wage adjustments (in particular in developing countries during their catch-up period); and the decline of energy demand due to carbon taxation triggers a significant drop of energy prices with respect to the baseline. Therefore, long-term costs are moderated[14].

3. A long distance race: technical change versus inertia

Let us now analyze in more detail how the above mechanisms work during the four phases of our time profiles.

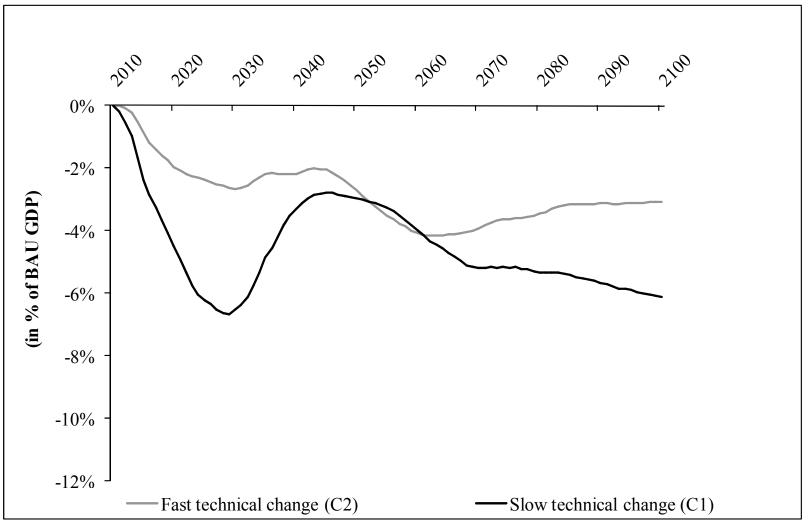
(i) During the 2010-2030 period, the particularly high GDP losses of climate policy found with IMACLIM-R are due to the sharp increase of carbon prices τ_E (Figure 3) and inertias in the decrease of *e/l*. Under adaptive expectations indeed, investment choices can be redirected only with high carbon prices whereas, under perfect foresight long-term prices are internalized in short term decisions which makes high short-term prices unnecessary. These carbon prices trigger increases of production costs, final prices and households' energy bills because the decrease of the carbon-intensity of the economy is limited by inertias on installed capital and on the renewal of households' end-use equipment (residential appliances, vehicles). These effects combine to undermine households' purchasing power, generate a drop in total final demand, a contraction of production, higher unemployment (under imperfect

[14] To demonstrate why those long-term costs can even be negative, it is necessary to represent the imperfect allocation of investments under baseline, which brings about considering a multisectoral model at the expense of analytical solvability.

labour markets) and an additional weakening of households' purchasing power through lower wages.

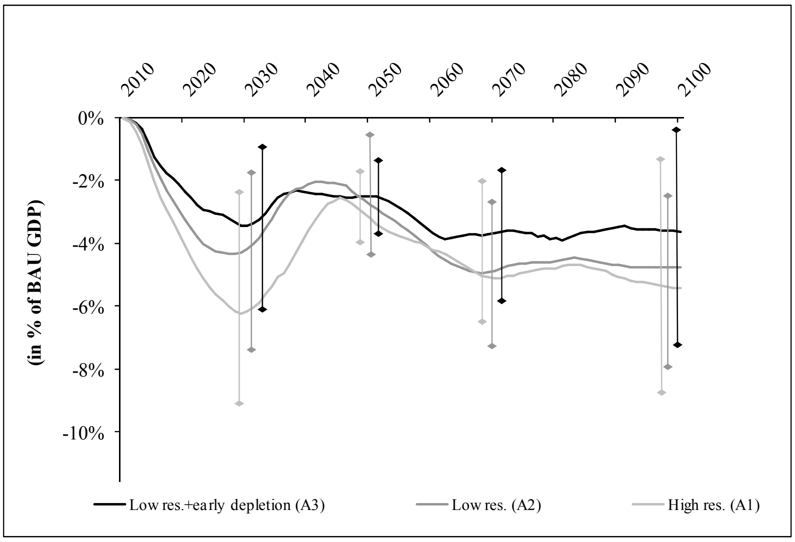
These mechanisms are more pronounced in emerging and developing countries because their industrial catch-up is based on a high share of energy-intensive basic industries with a high ratio *e/l* and low wages *w*. The GDP losses caused by these structural characteristics are enhanced by the negative effect of a unique carbon price on the international competitiveness of these carbon-intensive economies (Figure 3.4 and Table 3.2).

Figure 3.5: *Average GDP variations between stabilization and BAU scenarios with slow (black) and fast technical change (grey). Note: Vertical bars give the range of values across scenarios at some dates.*



Unsurprisingly, more optimistic assumptions on technological change limits short-term losses, since fast technical change partly counterbalances the inertia on the renewal of installed capital and makes decarbonisation easier: the energy intensity of production decreases, the carbon price necessary to trigger decarbonisation is lower, and those two effects combine to reduce GDP losses (Case C-2 in Figure 3.5). In addition, transitory costs are much lower where low oil reserves impose high short-term oil prices further accelerating technical change (Case A-3 in Figure 3.6).

Figure 3.6. *Average GDP variations between stabilization and BAU scenarios with the three assumptions on oil and gas supply. Note: Vertical bars give the range of values across scenarios at some dates.*



(ii) Between 2030 and 2045, the economic catch-up observed in Figure 2 is due to two major positive effects of early carbon prices, which lowers the weight of energy in the production process, $p_E.e$. First, moderation of oil demand in stabilization scenarios delays Peak Oil and the associated oil price increase (reduction of energy prices p_E). Second, the accumulation of learning-by-doing favours the diffusion of carbon-free technologies over this time horizon, with the co-benefit of enhanced energy efficiency (lower e). The mitigation costs are further moderated at this time horizon by the decrease of carbon price τ_E between 2030 and 2045, permitted by the abundance of mitigation potentials below 50\$/tC0$_2$ in the residential, industrial and power sectors (see (Barker et al., 2007, Figure TS27)). Those effects can be interpreted as a partial correction, *via* carbon pricing, of sub-optimal investment decisions in the BAU scenarios. A steady increase of fossil energy costs (carbon price included) partly compensates for the imperfect anticipation of increases in oil prices in the BAU scenario. It forces short-sighted decision-makers to progressively internalize constraints in fossil fuel availability, and accelerate the learning-by-doing in carbon-saving techniques. This yields a virtuous macroeconomic impact through a lower burden of imports in oil importing economies and reduced volatility of oil prices. In this sense, a carbon price is a hedging tool against the uncertainty on oil markets (Rozenberg et al., 2010).

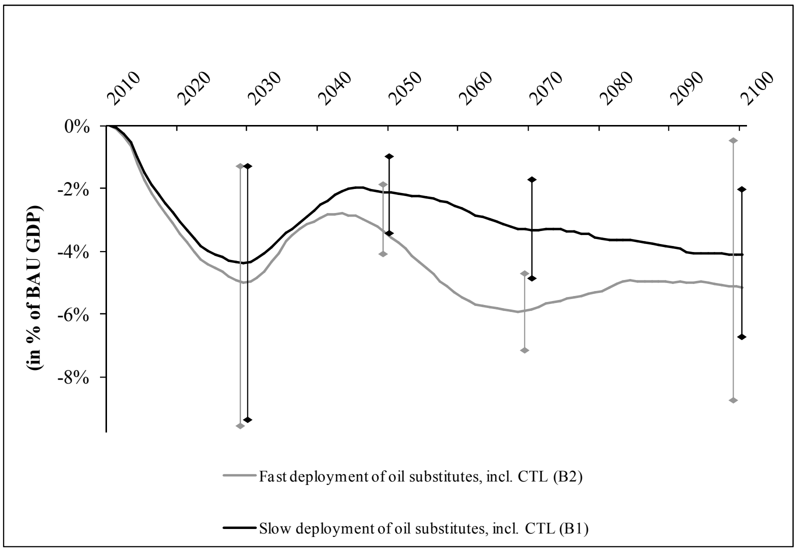
This virtuous effect is stronger in the case of slow technical change leading to high oil dependency (scenario C-1 in Figure 3.5). In this case indeed, economies are more vulnerable to Peak Oil in the BAU scenario. The GDP catch-up is also more important in the case of high reserves. In the Scenario A-1 (Figure 3.6) indeed, oil-free technical change is for long discouraged by low oil prices and the high economic burden imposed by Peak Oil period is significantly reduced by early carbon pricing.

(iii) Around 2050, a new phase of increasing mitigation costs starts as a consequence of a sharp increase of carbon prices τ_E from around 100\$/tCO2 in 2045 to around 300\$/tCO2 in 2070. Indeed, at this time horizon, most of the low cost mitigation potentials in the residential, industrial and power sectors have been exhausted, and the essential of emission reductions has to come from the transportation sector. A fast increase of carbon prices is then necessary to ensure emission reductions despite the weak sensitivity of the transportation sector to carbon prices and the trend of increasing carbon-intensive road-based mobility. This context is generated by the concomitance of four effects: a) the massive access to motorized mobility in developing countries, b) the absence of targeted policies to control urban sprawl, which tends to increase the dependence on constrained mobility c) the abundance of investments in road infrastructure, which decrease road congestion and favor the attractiveness of private cars at the expense of other transportation modes, d) the rebound effect on mobility demand consecutive to energy efficiency gains, which offsets approximately 25% of the emissions reductions that would have resulted from technical energy efficiency improvement.[15]

In addition the diffusion of Coal-To-Liquid (CTL) as a mature substitute to oil after 2050 makes passenger mobility particularly carbon intensive. During this post-Peak Oil period, the assumptions about the degree of maturation of CTL are one key determinant of overall costs. Mitigation costs are high with a rapid deployment of CTL in the BAU scenario, making very high carbon prices necessary to limit its penetration (Case B-2 in Figure 3.7)

[15] This order of magnitude of the rebound effect is in the range of empirical measures reported in the literature (Greening et al., 2000).

Figure 3.7. *Average GDP variations between stabilization and BAU scenarios with slow (black) and fast (grey) deployment of oil substitutes; Note: Vertical bars give the range of values across scenarios at some dates.*



(iv) After 2070, an increase of carbon price is still necessary to control emissions in the transportation sector since most other mitigation potentials have already been exploited (up to 600\$/tCO_2 in 2100). However, contrary to the first period, high carbon prices do not necessarily induce significant GDP losses. Indeed, they apply to a low-carbon economy and, at that time horizon, the share of labor costs has increased drastically in currently developing regions, making the critical energy-to-labor cost ratio $\left(\frac{p_E.e}{w_0 l}\right)$ far lower.

The race between increasing mobility needs and technical change in the transport sector is thus critical to explain discrepancies across scenarios over the century. In this race, the diffusion of Electric Vehicles is a key parameter (given almost carbon-free power generation). Optimistic assumptions for the market potential of electric vehicles accelerate its diffusion, decrease the energy cost $p_E.e$ and allows for a final phase of GDP catch-up (C-2 in Figure 3.5). Conversely, if the diffusion of electric vehicles is limited, the transportation sector remains fossil fuel intensive and further emission reductions come at as slightly increasing cost (C-1 in Figure 3.5).

III. Beyond carbon pricing: the role of investments in long-lived infrastructure

The time profile of mitigation costs obtained with a worldwide carbon price in Section 3 highlights two major concerns. First, the necessity of high short-term carbon prices is detrimental to most economies and triggers high transitory losses (especially in developing countries). The sensitivity of short-term mitigation costs to technological change parameters, as discussed in Section 3.3(i), suggests that support schemes to low-carbon technologies may be an appropriate complementary measure to foster early investments and endogenous improvements of low-carbon technologies (Kverndokk and Rosendahl, 2007) and to lower carbon prices (see Bosetti et al., 2009).[16]

Second, these high short-term carbon prices may be insufficient to limit long-term losses, notably because of the very specific dynamics of the transportation sector (Jaccard, 1997) where energy prices are swamped by other determinants (e.g., real estate markets, political bargaining behind infrastructure policies and just-in-time processes in the industry). In this section, we test a design of climate policy where carbon pricing is complemented by measures aimed at controlling the long-term dynamics of transport-related emissions. In this exploratory exercise, we represent spatial planning-related policies and changes in investment decisions for long-lived infrastructure in a synthetic way through three main sets on assumptions[17].

(i) a shift in the modal structure of investment in transportation infrastructure favoring public modes against private cars. Instead of assuming that the allocation of investments follows modal mobility demand, we consider public policies for reallocating part of them from road to low-carbon transportation infrastructure (rail and water for freight transport, rail and non-motorized modes for passenger transport).

[16] These schemes are not investigated explicitly in this paper, but are implicitly captured by the assumption on technological change.

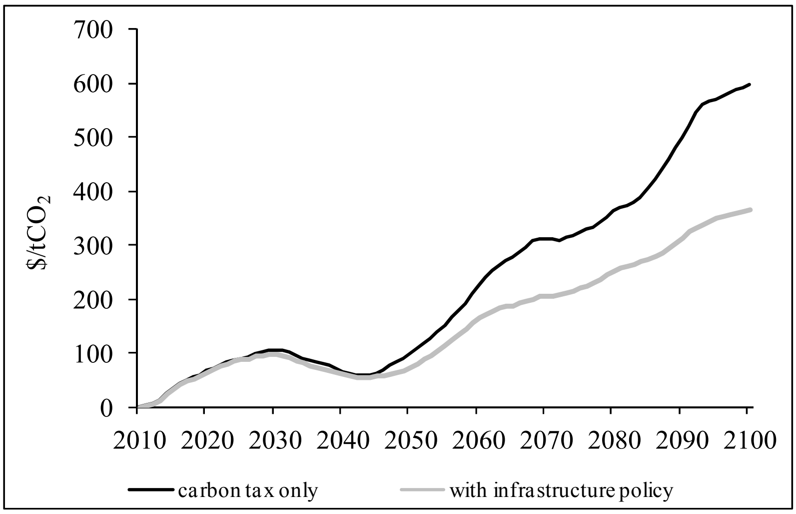
[17] Given the absence of reliable and comprehensive data on the cost of implementation of these measures, we assume a redirection of investments at constant total amount and neglect side costs and benefits.

(ii) a progressive relocation of buildings infrastructure that allows for a reduction of households' constrained mobility (essentially commuting) from the 50% of total mobility as previously considered to 40% .

(iii) changes in the production/distribution processes allowing to reduce transport needs (we considered a 1% decrease of the input-output coefficient between transport and industry to be compared with a constant coefficient in the previous case).

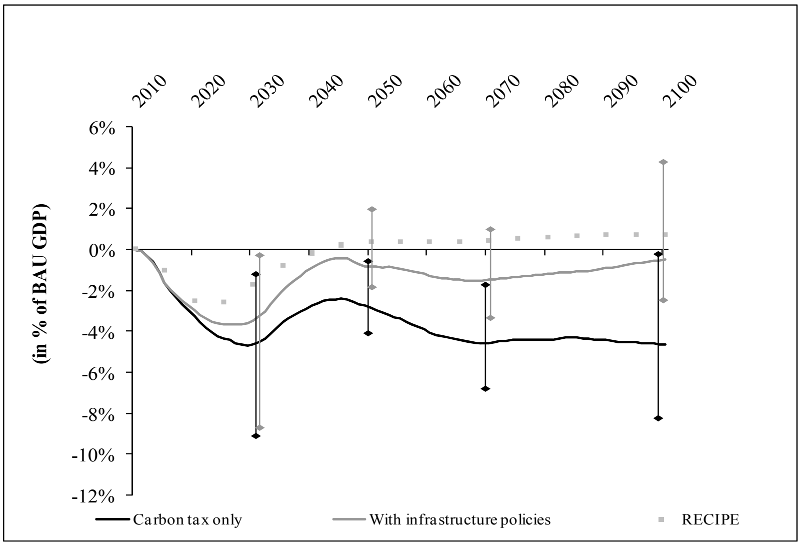
Replicating the numerical experiments for the 12 above BAU scenarios, we find that the reduction of mobility needs and the shift towards low-carbon modes allows meeting the same climate objective without a steep rise of carbon prices over the long run (Figure 3.8) and with far more moderate GDP losses (Figure 3.9).

Figure 3.8: *Average carbon price for a 'carbon price-only' policy (black) or with complementary infrastructure policies (grey).*



Another important finding is that these positive effects become especially important only after 2050, which is the logical outcome of the inertias in deploying new infrastructures. However, the complementary measures do not change drastically carbon prices before 2050, their impact on GDP losses is already visible between 2025 and 2050 in the form of an acceleration of the GDP catch-up with its BAU level and negligible losses around 2050 (Figure 9). The alternative infrastructures are not fully deployed but they begin to have an influence on the demand for gasoline (constraining the rebound effect) at the very moment when other sources of decarbonization start to become exhausted. This reduction of gasoline demand has a significant impact on the dynamics of the oil market and yields a deeper 'peak oil avoidance.

Figure 3.9: *Average GDP variations between stabilisation and BAU scenarios for a 'carbon price only' policy (black) or with accompanying infrastructure policies (grey). Note: Vertical bars give the range of values across scenarios at some dates; squares represent GDP variations between the stabilisation and the BAU scenario in the RECIPE project.*



In the short run, the effects of complementary policies is less important, but non negligible; as GDP losses being reduced by 20% in 2025 with respect to the carbon-price only policy. But, because of the inertia of transportation infrastructures, the bulk of the transition problem has to be addressed through other policies (fiscal policies, differentiated tariffs, subsidies for energy efficiency in the residential sector, etc.) targeted to avoid a full transmission of the carbon price to householders' energy bill, especially in developing countries.

IV. Conclusion

This paper analyzes the macroeconomic effects of a world carbon price with the IMACLIM-R model. The profiles of GDP costs differ significantly from those found in a first best economy, because the model captures key features of second- best economies (non fully flexible labor markets, imperfect competition, adaptive foresight) and represents the inertia of technical systems. We demonstrate the role of infrastructure dynamics in the formation of these costs profiles and investigate a richer climate architecture, where carbon pricing is complemented by measures designed to control transport-related emissions.

Over the short term (2020-2030), the absence of perfect foresight makes high carbon prices necessary and causes high GDP losses, especially in developing countries. Technical inertia limits the pace of decarbonization and the high carbon prices increase production costs; these costs are transmitted to the selling prices and combine with higher household's energy bill to undermine consumers' purchasing power. This is more important in emerging economies because of their higher labor-to-energy costs ratio.

In the medium term (2030-2050) the reduced vulnerability to Peak Oil and the acceleration of learning-by-doing in fossil-free techniques allow the policy trajectories to catch-up the BAU GDPs. Over the long-term (2050-2100), there is the recurrence of significant costs. Indeed, the high short-term carbon prices prove to be insufficient to shift the patterns of urban and transportation infrastructures and prevent an explosion of road based mobility. In the absence of very optimistic assumptions on biofuels and electrical cars, high carbon prices are necessary to outweigh these trends in the long-term period where transport represents the core of emissions reductions.

This lock-in can be avoided by specific measures triggering an early redirection of investments in favor of modal shifts towards public modes, moderation of urban sprawl, and curtailment of the transport intensity of production. The adoption of such measures proves to significantly reduce the policy costs in the short and medium term and to create room for low and even negative long-term costs. This analysis highlights the importance of measures designed to shift investments in long-lived infrastructures as complementary policies to carbon pricing. More detailed insights could be obtained by disaggregating these measures in a set of agglomeration-specific policy measures, and complementary initiatives at alternative spatial scales, that may shed light on different behavioral responses, in terms of relocation of production and consumption activities.

Appendix: Numerical assumptions and variants of scenarios

A. Numerical assumption on oil and gas supply

The three crucial determinants of the 'oil supply' Nexus are the amount of ultimate resources (and their regional distribution), the inertia on capacity deployment and the decision of Middle-East producers acting as "swing producers".

Most estimates of proved oil reserves converge around 2.2 Tbbl (BP, 2011) including past production. To reflect controversies about the amount of reserves to be discovered, we adopt two assumptions for ultimate resources Q_∞: 3.3 Tbbl and 3.8 Tbbl. The lower bound reflects a conservative assumption on resource additions, in line with estimates from the Association for the Study of Peak Oil (ASPO). The higher bound considers higher resource potentials, corresponding to median estimates by (USGS, 2000; Greene et al., 2006; Rogner, 1997).

The intensity of constraints on production growth due to geological constraints is captured by the slope parameter b[18]. For conventional oil, we adopt the econometric estimate from Rehrl and Friedrich (2006): b_C=0.061/year. Given uncertainty on large scale production of non-conventional oil, we consider either the same value than conventional oil, b_{NC}=0.061/year, or more pessimistic assumption of a slower deployment with b_{NC}=0.04/year. For Middle-East producers, we impose in addition a cap on the annual increase of production capacity, ΔCap_{ME}.

The deployment of production capacities in Middle-East countries is decided by the price objective p_{obj}. A benchmark for oil price setting is a continuous increase towards a medium-term stabilization around 80$/bbl, reflecting the progressive loss of influence of Middle-East producers. Given uncertainties, especially in the geopolitical context, we also consider the possibility that Middle-East producers are able to expand their production capacities to bring oil price at their pre-2004 level, 40$/bbl. This market flooding option is possible only for the more optimistic assumption on reserves. This exercise of the market power ends up when the finiteness of the resource forces a decline of production. For the sake of simplicity, we assume that it happens once a share sh_D of their reserves remains underground, and consider two values (50% or 25%) to reflect the uncertainties on the stock of resource in Middle-East countries.

[18] a small (high) b means a flat (sloping) production profile to represent slow (fast) deployment of production capacities.

Finally, the 'gas supply' NEXUS represents indexation of gas markets on oil markets with a 0.68 elasticity of gas to oil price, as calibrated on the World Energy Model (IEA, 2007). But, in order to represent the possibility that gas scarcity triggers faster price increases, we consider an alternative where this indexation disappears when oil prices exceed a threshold level $p_{oil/gas}$ (chosen at 80$/bbl). In this latter case, gas prices are driven by the increased margins for gas producers.

These numerical assumptions are grouped in three variants summarized in Table 3.3

Table 3.3: *Numerical assumptions for the three variants on oil and gas supply*

	Unit	**A-1**	**A-2**	**A-3**
Q_∞	*Tb*	3.8	3.3	3.3
p_{obj}	*$/bl*	40	80	80
b_{NC}	*Year^{-1}*	0.061	0.04	0.04
ΔCap	*Mbl/y*	0.8	0.7	0.7
sh_D	*%*	25	25	50
$p_{oil/gas}$	*$/bl*	80	∞	∞

B. Numerical assumption on substitutes to oil

The 'alternatives to oil' Nexus considers two large-scale substitutes to oil for liquid fuels production: biofuels and Coal-To-Liquid.

The supply curves, $S_{bio}(t,p)$ give biofuels production, given competition with oil, and are taken from IEA (2006). They assume maximum biofuels production at 14 EJ/year in 2030 and, thanks to technical progress, at 42 EJ/year in 2050. These assumptions are quite conservative with respect to recent estimates about biofuels potential (Chum et al, 2011, Figure 2.23(b)) and we introduce an alternative, more optimistic, assumption allowing 20 EJ/year in 2030 and 60EJ/year in 2050. The diffusion of biofuels is in addition submitted to the constraint of a time delay, Δt_{bio}, which captures inertia on the deployment of raw products (biomass) and of refining capacity.

Coal-To-Liquid is treated as a backstop technology, which enters the market as soon as liquid fuel selling price exceeds its total cost, p_{CTL}, including production processes and risk premium. This backstop technology is submitted to capacity constraints in the form of a delay Δt_{CTL} between investments and production. Given uncertainty on large-scale CTL production, we consider two possibilities, depending whether CTL is a mature technology (low threshold oil price at 120$/bbl and no inertia in the deployment) or it is submitted to constraints slowing down its deployment (high threshold oil price at 200$/bbl and significant time-lag in the deployment).

These numerical assumptions are grouped in two variants summarized in Table 3.4

Table 3.4: *Numerical assumptions for the two variants on oil substitutes*

	Unit	**B-1**	**B-2**
$S_{bio}(t,p)$	*Mtoe/y*	$\underline{S_{bio}}(t,p)^{(*)}$,	1.5* $\underline{S_{bio}}$ (t,p)
Δt_{bio}	*Years*	6	4
p_{CTL}	*$/bl*	200	120
Δt_{CTL}	*Years*	8	0

(*) *exogenous trend from (IEA, 2006)*

C. Numerical assumptions on demand-side technical change

The ‘Power generation’ Nexus represents investment choices in new power generation technologies according to a minimization of mean production costs. Technical change is then dependent upon the decrease of capital costs, along with the learning process controlled by technology-specific learning rates γ (it measures the percentage decrease of capital costs for each doubling of experience). Learning does not affect standard technologies due to saturation of experience, but potentially contributes to important costs decreases in more recent or prospective technologies, including wind energy and Carbon Capture and Storage. Due to uncertainties on the technical potentials of these technologies, we represent either fast learning through high learning rates (7% for wind *vs* 13% for CCS) or constrained learning with low learning rates (3% for wind *vs* 7% for CCS). Note that we consider lower learning rates for wind units than for CCS to represent that the former is a more mature technology, with less remaining progress potential.

In addition, the ‘Power generation’ Nexus represents the constraints that may affect the diffusion of carbon-free power plants by an exogenous maximum market share, with different dynamics for already existing and new technologies. In the former group, we explicitly represent Nuclear and Wind Energy and assume their maximum shares Sh_{Nuke} and Sh_W as constant-over-time. We adopt rather conservative assumptions on the long-term potential of Nuclear and consider a maximum market share at 40% to capture limitations for social acceptability reason (20% in a more constrained vision). For wind energy, we consider a

benchmark case where it is limited to 15% of production to capture implicitly constraints imposed by intermittent production and additional integration costs at higher shares. This assumption is in line with the median estimate of the 164 global scenarios reviewed by the IPCC (Wiser et al, 2011, Figure 7.25). But, a growing body of work has evaluated higher levels of deployment, around 20% or more, provided that cost and policy factors are favourable. To treat this case, we also consider a higher limit on wind's market share, at 25%. In the latter group, we consider Carbon Capture and Storage (CCS), and the maximum share Sh_{CCS} increases over time to represent its progressive deployment, ranging from zero at the starting year ($t_{0,CCS}$) up to its long-term market potential $Sh_{max,CCS}$. During the early years, inertia limits the deployment of this new technology as captured by a slow increase of the maximum share during a 'bottleneck period' of length $\Delta t_{CCS,}$ followed by an accelerated increase once the technology is mature.
In the 'Industry and services' Nexus, energy prices affect the selection of new production capacities but do not influence existing ones. This putty-clay assumption implies that changes in final energy use are dependent on their lifetime Δtind. This is an important variable, since it conditions both the turnover of productive capital (and hence the speed of technical change) and investments needs. We take 20 years as a benchmark case, whereas 30 years reflects a more constrained context on investment imposing delayed retirement of production capacities.
In the 'Housing and Buildings' Nexus, the baseline trends of energy consumption per square meter, $\alpha_{res}(t)$ are taken from outcomes of the POLES model. They feature a relative decrease of unitary energy demand in developed regions thanks to energy efficiency, while strong increases in developing countries are due to the access to energy services along with wealth increase. In addition, the energy mix is orientated towards electricity and gas at the expense of coal and oil. We consider also more energy-intensive pathways with proportionally higher unitary energy consumption due to lower efficiency gains (for technical constraints or lack of investments) and/or a more prominent access to energy services in developing countries.

In the 'Freight transportation' Nexus, the energy intensity of vehicles is driven by an exogenous trend $\mu_f(t)$ and a short-term fuel price elasticity ε_f to capture autonomous and endogenous energy efficiency gains as well as short-term modal shifts, respectively. The long-term price response of the fleet then results from the sequence of those short-term adjustments.

The 'Passenger Transportation' Nexus represents the crucial determinant of energy efficiency and modal choices. Energy efficiency in private transportation is mainly dependent on the constraints on the diffusion of Electric Vehicles (EV). They are captured by an exogenous maximum *Sh* on their market share, which ranges from zero in the first year ($t_{0,EV}$) to $Sh_{max,EV}$ as it achieves its long-term market potential. During the early years, inertia limits the deployment of this new technology, as captured by a slow increase in the maximum share during a 'bottleneck period' of length $\Delta t_{EV,}$ followed by an accelerated increase once the technology has matured.

Modal allocation of mobility demand is affected by investments in infrastructure, which determine the relative efficiency of the different modes. Instead of the default assumption that investment is allocated proportionally to modal mobility demand, alternative decisions may trigger a re-allocation from road to low-carbon transportation infrastructure (public and rail transport for passengers and rail and water transport for freight).

These numerical assumptions are grouped in two variants summarized in Table 3.5.

Table 3.5: *Numerical assumptions for the three variants on delmand-side technical change*

		Unit	**C-1**	**C-2**
Nuclear	Sh_{Nuke}	%	20	40
Wind energy	Sh_W	%	15	25
	γ_W	%	3	7
Carbon Capture and Storage	$t_{0,CCS}$	*Date*	2015	2010
	Δt_{CCS}	*Years*	10	7
	$Sh_{max,CCS}$	%	70	100
	γ_{CCS}	%	7	13
Electric Vehicles	$t_{0,EV}$	*Date*	2020	2010
	Δt_{EV}	*Years*	8	3
	$Sh_{max,EV}$	%	50	80
	γ_{EV}	%	10	20
Freight transport	$\mu_f(t)$	-	1	$\underline{\mu_f}(t)^{(**)}$
	ε_f	-	-0.35	-0.4
Buildings	$\alpha_{res}(t)$	*Toe/m²*	$1.2* \underline{\alpha_{res}}(t)$	$\underline{\alpha_{res}}(t)^{(**)}$
Industry	Δt_{ind}	*Years*	30	20

$^{(**)}$*exogenous trend from the POLES energy sectoral model(LEPII-EPE ,2006)*

REFERENCES

Armington PS (1969). A Theory of Demand for Products Distinguished by Place of Production. *IMF, International Monetary Fund Staff Papers* 16: 170-201.

Arrow KJ and Debreu G (1954). Existence of an equilibrium for a competitive economy. Econometrica 22: 265-290.

Barker T, Bashmakov I, Bernstein L, Bogner JE, Bosch PR, Dave R, Davidson OR, Fisher BS, Gupta S, Halsnæs K, Heij GJ, Kahn Ribeiro S, Kobayashi S, Levine MD, Martino DL, Masera O, Metz B, Meyer LA, Nabuurs G-J, Najam A, Nakicenovic N, Rogner HH, Roy J, Sathaye J, Schock R, Shukla P, Sims REH, Smith P, Tirpak DA, Urge-Vorsatz D, Zhou D (2007): Technical Summary. In: *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [B. Metz, O. R. Davidson, P. R. Bosch, R. Dave, L. A. Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Barker T and Scrieciu SS (2010). Low stabilization within a "New Economics" macro-econometric framework: insights from E3MG. In: Edenhofer, O., Knopf, B., Leimbach, M., Bauer,N. (eds) The Economics of Low Stabilization, The Energy Journal 31(Special Issue 1), in press.

Bosetti V, Carraro C, Massetti E, Sgobbi A and Tavoni M (2009). Optimal Energy Investment and R&D Strategies to Stabilise Greenhouse Gas Atmospheric Concentrations, Resource and Energy Economics 31(2): 123-137.

BP (2011). BP Statistical Review of World Energy, June 2011. available at: http://www.bp.com/assets/bp_internet/globalbp/globalbp_uk_english/reports_and_publications/statistical_energy_review_2011/STAGING/local_assets/pdf/statistical_review_of_world_energy_full_report_2011.pdf

Chum H, Faaij A, Moreira J, Berndes G, Dhamija P, Dong H, Gabrielle B, Goss Eng A, Lucht W, Mapako M, Masera Cerutti O, McIntyre T, Minowa T and Pingoud K (2011). Bioenergy. In: O. Edenhofer, R. Pichs-Madruga, Y. Sokona, K. Seyboth, P. Matschoss, S. Kadner,, T. Zwickel, P. Eickemeier, G. Hansen, S. Schlömer, C. von Stechow (eds) IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation , CambridgeUniversity Press, Cambridge, United Kingdom and New York, NY, USA.

Corrado C and Mattey J (1997). Capacity Utilization. Journal of Economic Perspectives 11(1): 151-167.

Edenhofer O, Carraro C and Hourcade JC (2010a). The rationale of modeling comparison exercises for climate policy. *accepted for publication in Climatic Change.*

Edenhofer O, Knopf B, Barker T, Baumstark L, Bellevrat E, Chateau B, Criqui P, Isaac M, Kitous A, Kypreos S, Leimbach M, Lessmann K, Magné B, Scrieciu S, Turton H, and van Vuuren D (2010b). The Economics of Low Stabilization: Model Comparison of Mitigation Strategies and Costs. In: Edenhofer, O., Knopf, B., Leimbach, M., Bauer,N. (Eds) The Economics of Low Stabilization. *The Energy Journal* 31 (Special Issue 1):11-48.

Grazi F, van den Bergh JCJM and van Ommeren JN (2008). An Empirical Analysis of Urban Form, Transport, and Global Warming. The Energy Journal 29(4), 97-107.

Greene D, Hopson J and Li J (2006). Have we run out of oil yet? Oil peaking analysis from an optimist's perspective. Energy Policy 34(5), 515-531.

Greening L, Greene D and Difiglio C (2000). Energy efficiency and consumption – the rebound effect – a survey. Energy Policy 28(6):389-401.

Guivarch C, Crassous R, Sassi O and Hallegatte S (2010). The costs of climate policies in a second best world with labour market imperfections, *forthcoming in Climate Policy.*

Hourcade JC, Jaccard M, Bataille C et Ghersi F (2006). Hybrid Modeling : New Answers to Old Challenges. *In:* Hybrid Modeling of Energy-Environment Policies: reconciling Bottom-up and Top-down. The Energy Journal (Special Issue 2): 1-12.

IEA (2007). World Energy Outlook 2007. *IEA/OECD*, Paris, France.

IEA (2006). Energy Technology Perspectives: Scenarios and Strategies to 2050. O*ECD/IEA*, Paris, France.

IPCC (2007). Summary for Policymakers. In: B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds) Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. CambridgeUniversity Press, Cambridge, United Kingdom and New York, NY, USA.

Jaccard M, Failing L and BerryT (1997). From equipment to infrastructure: community energy management and greenhouse gas emission reduction. Energy Policy 25(13): 1065- 1074.

Jakob M., Luderer G, Steckel J, Bosetti V, Tavoni M, and Waisman H. (2010). Time to act now? Assessing the costs of delaying climate measures and benefits of early action, *accepted for publication in Climatic Change*.

Johansen L (1959). Substitution versus Fixed Production Coefficients in the Theory of Growth: A synthesis. Econometrica 27: 157-176.

Kverndokk S and Rosendahl KE (2007). Climate Policies and Learning By Doing: Impacts and Timing of Technology Subsidies. Resource and Energy Economics 29: 58-82.

LEPII-EPE (2006). The POLES model, *Institut d'Économie et de Politique de l'Énergie*, Grenoble, France, 12 pp. Accessible at : http://webu2.upmfgrenoble.fr/iepe/textes/POLES12pJan06.pdf.

Luderer G, Bosetti V, Jakob M, Steckel J, Waisman H and Edenhofer O (2010a). The Economics of GHG Emissions Reductions – results and insights from the RECIPE model intercomparison. *accepted for publication in Climatic Change*.

Luderer G, Leimbach M, Edenhofer O and Hourcade JC (2010b). The regional distribution of mitigation costs – a tale of scarcity rents . *accepted for publication in Climatic Change*.

Malcolm G and Truong P (1999). The Process of Incorporating Energy Data into GTAP, *Draft GTAP Technical Paper*, Center for Global Trade Analysis, Purdue University, WestLafayette, Indiana, USA.

Paltsev S, Reilly J, Jacoby H, Eckaux R, McFarland J, Sarofim M, Asasoorian M, Babiker M (2005). The MIT Emissions Prediction and Policy Analysis (EPPA) Model: Version 4, Report no. 125. *Joint Program on the Science and Policy of Global Change*, MIT, Cambridge, MA

Rehrl T and Friedrich R (2006). Modeling long-term oil price and extraction with a Hubbert approach: The LOPEX model. Energy Policy 34(15): 2413-2428.

Rogner HH (1997). An assessment of world hydrocarbon resources. Annual Review of Energy and the Environment 22:217–262.

Rozenberg J, Hallegatte S, Vogt-Schilb A, Sassi O, Guivarch C, Waisman H and Hourcade JC. (2010). Climate policies as a hedge against the uncertainty on future oil supply, *Climatic Change (forthcoming)*

Sands RD, Miller S and Kim MK (2005). The Second Generation Model: Comparison of SGM and GTAP Approaches to Data Development, *Pacific Northwest National Labouratory*, PNNL-15467, 2005.

Sassi O, Crassous R, Hourcade JC, Gitz V, Waisman H and Guivarch C (2010). IMACLIM-R: a modeling framework to simulate sustainable development pathways. Int. J. Global Environmental Issues *10(1/2)*:*5–24*.

Solow R (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70(1): 65-94.

United States Geological Survey (USGS) (2000). World Petroleum Assessment 2000. USGS, Washington.

Wiser R, Yang Z, Hand M, Hohmeyer O, Infield D, Jensen PH, Nikolaev V, O'Malley M, Sinden G, Zervos A (2011). Wind Energy. In: O. Edenhofer, R. Pichs-Madruga, Y. Sokona, K. Seyboth, P. Matschoss, S.Kadner, T. Zwickel, P. Eickemeier, G. Hansen, S. Schlömer, C. von Stechow (eds) IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation,Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Part B

Urban land and the spatial dimension of climate policies

This introductory section builds upon the general theory of rents (see General introduction) to identify the crucial determinants of urban land markets.

I- Spatial differentiation and differential production rents

The role of the spatial dimension of economic activities has received much less attention than the temporal dimension in the standard theory of rent. Such approaches ignoring the spatial dimensions consider implicitly that the economic mechanisms driving the emergence of rents are independent of land-use patterns. This assumption is questionable when introducing an immobile factor, since location becomes then a discriminating factor likely to give a comparative to those producers benefiting from the more interesting locations. This dimension has been introduced by Von Thünen (1836), who studies equally fertile lands used for agricultural production in the case where the goods can be sold in a single, immobile point of space. The selling price will be identical independently from the location where it is produced, but the production costs depend on the distance between production and market locations, as a result of transport costs. In this case, producers located in the central location will benefit from a surplus with respect to further locations, which gives rise to differential rents in the similar manner than in the standard Ricardian theory.

This idea that location can be an important determinant of rents is extended to non-agricultural production by Chamberlin (1933), who considers retail sales in urban areas where consumers' moving are constrained. This immobility factor introduces a spatial differentiation *via* the size of the market at producers' disposal, which results in imperfect competition, differentiated prices and ultimately different rent levels in function of the location. the excess of production obtained by firms on the better locations is a form of a differential rent, resulting from better access to the market.

II- From production rents to land rents

The share of production rents that is transferred to the landowner depends on the natire of land competition :

> « The differential remaining, which is due to the superiority of the profit making opportunities afforded by one site as compared to another, is rent, and is put into the

hands of landlords by the competition of entrepreneurs for the best opportunities» (Chamberlin, 1933)

The nature of land rents perceived by landowners in this situation is subjected to theoretical debates (Evans, 1991; Foldvary, 1993; Evans, 1993). In the extreme case where a unique agent owns all land, he can freely differentiate them according to production surpluses permitted by each location characteristics, and can then capture the whole production rents in the form of land revenues paid by producers. In this case, even if the landowner can capture the rent thanks to his monopoly power on land markets, land rents are differential in nature since they pay the advantages of the location for production. In more realistic cases where several landowners are active, land competition prevents them from deciding the price of their plot of land independently from others, and the landowner only captures a share of total production rents perceived by producers. Land price at a given location is then determined by the competition between different uses of a given plot and the share of land rent perceived by landowners then depends on its scarcity, given its intrinsic characteristics,

III- The specific case of residential uses

In this thesis, we focus more specifically on land use in urban environment for households' housing purposes, which enters in the more general framework described above when considering housing services as the good produced from land use. Housing prices then basically obey to differential principles, the share of which that is captured by landowners being dependent on the scarcity of the site under consideration. The first consistent depiction of these mechanisms is provided by Alonso (1964), who generalizes Von Thünen's model to non-agricultural urban agents. He investigates location choices of three categories of agents (farmers, industrial and households) with respect to the Central Business District (CBD) where all economic activities are concentrated on the basis of a tradeoff between transport costs and housing costs. In this approach, the analysis of urban rents is based on the bid-rent function, which measures what agents are ready to pay to ensure the possession of a given location. The equilibrium spatial distribution of economic activities then results from competition between the different categories of agents, the more profitable activities occupying the better locations (i.e. the closest to the CBD) because they can afford higher land prices. Muth (1969) extends Alonso's approach to make it suitable for the question of residential locations by introducing core determinants of the desirability of a given location within a city beyond the distance from the CBD, namely construction constraints and broader dimensions of the urban environment affecting households' utility. In this framework, housing rents are continuously increasing when getting closer from CBD, since transport costs weight less on households' budgets and allows them to pay higher land rents. Housing rents then appear as the difference between this bid-rent price, decided according to competition for better residential location, and construction costs.

IV- Land rent and monopoly power

A monopoly rent can emerge over the above described sources of rent in the case where landowners can act as monopolist by reducing the supply of housing services provided on their plot of land through limited construction. In this case, enhanced competition fosters a rise of prices and then of rents.

V- Specifications for the representation of urban land price

The analysis of mechanisms driving the appearance and amount of land rents in a urban environment gives us the fundamentals of the formation of prices for residential uses:

- Land prices in urban environment are driven by the competition between residential and other production uses, and the price paid by households results from the share of production rents that can be captured by landowners. In a simplified vision where housing builders (who invest capital on land to provide housing services) are landowners, this rent transfer is total and land price effectively represents the price paid by households for residential services. This assumption is not a strong limitation for this thesis, which is interested in the economic effects of land prices on households' budgets.[1]
- Land price results from the superposition of differential rents reflecting the specificities of the locations in terms of accessibility and amenities, and monopoly rents allowed by a localized market power, which play an important role fro residential uses in reason of the constraints on construction.

[1] If we wanted to consider the effects of land taxation in the spirit of Henry George (1879), it would be necessary to distinguish explicitly landowners and housing services builders, and to describe the mechanisms driving the distribution of rents among those two categories of agents

Chapter 4 and 5 describe the theoretical foundations and the modeling assumptions adopted to capture these mechanisms of urban land markets. Chapter 6 finally enables a coupled version of IMACLIM, including the description of urban areas to provide a consistent vision of the interplay between energy, mobility and urban dimensions.

REFERENCES

Alonso WA (1964). Location and land use: toward a general theory of land rent. Harvard University Press, Cambridge.

Chamberlin, EH (1933). The theory of monopolistic competition. Harvard University Press. Cambridge

Evans AW (1993). “On Monopoly Rent: Reply”. Land Economics 69 (1): 111-112.

Evans AW (1991). “On Monopoly Rent”. Land Economics 67(1): 1-14.

Foldvary FE (1993). “On Monopoly Rent: Comment”. Land Economics 69(1): 108-110.

George H (1879). Progress and Poverty. New York: Robert Schalkenbach Foundation.

Hotelling H (1931). “The Economics of Exhaustible Resources”. Journal of Political Economy 39:137-175.

Muth RF (1969). Cities and housing. Chicago: University of Chicago Press.

Von Thunen JH (1826). Der Isolierte Staaat in Beziehung auf Landtschaft und Nationalokonomie, Hamburg. (English translation by C.M. Wartenburg: von Thunen's Isolated State, Oxford, Pergamon Press)

Chapter 4

The economic geography of environmental sustainability

Chapter 4[*] proposes a theoretical analysis of location decisions when long-term negative environmental externalities are accounted for. To this aim, we develop a theory of "spatial sustainability" by combining industrial location decisions, production- and trade-related environmental externalities, and the dynamics of migration and pollution. This study generalizes earlier modeling efforts to address agglomeration and environmental externalities in a location-trade framework. The model is innovative in that it: *i*) formalizes heterogeneous patterns of land development in multiple regions; *ii*) includes an endogenous agglomeration effect, which influences environmental pollution through two opposite mechanisms; *iii*) allows for continuous and asymmetric distributions of population and economic activities within the region; and *iv*) distinguishes between environmental externalities and sustainability.

[*] This chapter is the reproduction of : Grazi F, Waisman H, van den Bergh JCJM (2011). The economic geography of environmental sustainability, *submitted to Journal of Environmental Economics and Management*

Translating the notion of global sustainable development into concrete principles and actions at local, regional, and national levels has turned out to be difficult [OECD (2007; 2010)]. One reason is that there is no agreed framework for studying the spatial dimension of sustainable development. The relation between international trade and environment has received considerable attention, but most of the literature ignores dynamic issues related to sustainability (Copeland and Taylor, 2004).

This paper presents a theoretical framework for analyzing the impact of spatial structure of the economy on its long-run sustainability. We introduce the notion of "spatial sustainability" to denote that spatial configurations and economic dynamics are consistent with environmental constraints, as defined by the assimilative capacity of the environment (Pezzey and Toman, 2005). One may think here of global environmental issues, notably the emission of greenhouse gasses (GHG) due to energy use, which gives rise to climate change (IPCC, 2007).

We develop a general equilibrium model motivated by the new economic geography (NEG) (Krugman, 1991). It integrates three important economic mechanisms which influence the (un)sustainability of the economy: namely, agglomeration externalities; the advantages of international or interregional trade; and the dynamics of migration and pollution. Agglomeration externalities, such as shorter travel distances and technological spillovers and knowledge sharing, affect the emission of pollution by manufacturing firms through their impact on the efficiency of transport-related energy inputs in production. To address the direct and indirect energy-use effects of spatial economic organization, the model includes an intermediate intra-regional transport sector. The model is used to study the long-run performance of the economy in terms of sustainability, and the emergence of alternative spatial configurations of the global economy under different parameter values.

The model extends the standard NEG framework in four main ways. First, it explicitly accounts for spatial structure through the design of various spatial configurations of the economy (urbanized versus undeveloped) and transition between these. Second, it includes a distance-related endogenous agglomeration effect, which allows to simultaneously model increasing returns to scale at the firm level and external economies at the industry level. Third, it can handle continuous and asymmetric distributions of population and economic activities within the region which is consistent with infinite set of trade costs. This in turn allows for realistic application of the model findings to environmental policy analysis. Fourth, the model includes the interplay of dynamic mechanisms of pollution and agglomeration to formalize the notion of spatial sustainability of the economy in relation to standard pollution dynamics and “nature’s regenerative capacity” (Copeland and Taylor, 2004), or its ‘pollution assimilation potential’. This ultimately allows distinguish between environmental externalities associated with domestic production and international trade at each point in ‘time’, and unsustainability due to cumulative pollution effects.

Following Krugman’s seminal work (1991), a considerable literature on the NEG has developed that addresses the mechanisms through which economies develop in space. It combines location choice, transport cost, trade barriers, and imperfect market competition in a mathematically tractable framework. Studies in this vein have addressed a variety of issues, including trade taxes, regulation of transport, and lobbying on factor mobility.[1]

Yet few studies that employ the NEG framework have covered environmental issues, and none has explicitly addressed its connection with spatial structure and sustainability. An early study, Brakman, Garretsen, Gigengack, van Marrewijk and Wagenvoort (1996), examined congestion as a dampening agglomerative force, but did not offer analytical solutions. Eppink and Withagen (2009) study biodiversity conservation in the context of regional economic specialization and development with an analytically solvable NEG model, but treat the environmental externality (i.e. biodiversity loss) as purely local. Rauscher (2003) develops a NEG model with pollution and obtains analytical solutions at the cost of assuming quasi-linear preferences, which gives a partial-equilibrium flavor to this approach. Zeng and Zhao (2006) investigate the “pollution haven” hypothesis[2] by embedding pollution into the standard “footloose capital” model, a variant of Krugman’s (1991) model that describes the migration of capital when labor is immobile (Martin and Rogers, 1995; Baldwin et al, 2003). However, the analytical tractability of this model is realized by ignoring the negative impacts of pollution on household utility.

In this paper we use an analytically solvable variant of Krugman’s (1991) model, the “footloose entrepreneur (FE)” model developed by Forslid and Ottaviano (2003). This model has become quite popular because it yields closed-form solutions. A disadvantage of it is that it may give rise to spatial equilibria that are more extreme than what one tends to find in reality (what is called the ‘catastrophic agglomeration’ result). As a consequence, applications of such a framework to the policy domain have been hampered by the unrealism of model outcomes (Ottaviano, 2003). By formalizing a smooth transition from economic

---

[1] For an overview of the NEG literature, see Fujita, Krugman and Venables (1999), Fujita and Thisse (2002), and Ottaviano and Thisse (2004).

[2] This hypothesis states that pollution-intensive industries will tend to move to countries with relatively lax environmental regulations. Many studies have performed an empirical test. For a survey and meta-analysis see Jeppesen, Folmer and List (2002).

agglomeration to dispersion and continuously variable degrees of heterogeneity of land development within each region, we overcome this limitation of the framework. This results in an approach that is applicable to policy questions.

A few other contributions have employed the Forslid and Ottaviano framework to incorporate pollution and analyze how this relates to agglomeration. Van Marrewijk (2005) and Grazi, van den Bergh and Rietveld (2007) focus on quasi-static and static short-run equilibria, respectively. Lange and Quaas (2007) provide a dynamic analysis of pollution and agglomeration, but do not consider the positive effects of agglomeration on pollution and sustainability through technological and knowledge-sharing spillovers. The result is a partial description of reality, resulting in environmental externalities that dominate the final equilibrium outcome in certain cases. In contrast, our model accounts for two effects of agglomeration on environmental pollution which work in opposite directions, as explained below.

Agglomeration spillover effects have received attention in the economic literature on trade theory and urban economics since Marshall and Chamberlin. Nevertheless, their formal representation has turned out to be difficult and controversial (Ciccone, 2002). Moreover, to the best of our knowledge, to date no study has achieved a simultaneous modeling of increasing returns to scale at the firm level à la Dixit and Stiglitz (1977), as is standard in NEG models, and agglomeration externalities at the industry level à la Scitovsky (1954). Our work tries to accomplish this, and for this purpose adds to the increasing returns to scale operating at the firm level an endogenous agglomeration effect variable defined at the industry level. In particular, the intensity of the agglomeration spillover effect is defined as a function of regional 'market density', captured by the number of firms that are active in the industry, and a regional 'market form' which is captured by the capacity of the infrastructure endowment.

Our approach also has a number of minor innovative features. First, since the focus of this paper is on spatial sustainability, we provide a detailed analysis of the effect of the parameters that relate to the spatial and environmental dimensions of the model, notably the degree of economic concentration and the intensity of the environmental externality for different spatial configurations. Second and related to this, the spatial dimension of the economy is strengthened through the introduction of an explicit domestic transportation sector, which contributes to intra-industry agglomeration externalities. Third, agglomeration affects environmental pollution through two mechanisms, which, ex ante make the net effect non-obvious. One effect is that agglomeration increases the scale of production activity by lowering the costs of production and hence leads to more energy use and associated emissions. The other is that agglomeration reduces the energy requirements for production through a lower intra-regional transport intensity of production (shorter distances) and technological (R&D and learning) spillovers. This in turn leads to an improvement in the energy-efficiency of technologies used by economic production activities and associated lower pollutive emissions.

The remainder of this paper is organized as follows. Section II develops a basic model. Section III derives the conditions under which qualitatively different, stable equilibria (with partial and full agglomeration) arise when agglomeration effects and pollution externalities are present. Section IV extends the model with the dynamics of pollution to derive long-run spatial equilibria that satisfy environmental sustainability and takes the next step of evaluating the sustainability performance of three different spatial configurations of the economy, resulting in a ranking of sustainable configurations. Section V concludes. Proofs of the main results are given in the Appendices.

I- Economy and Space

1. The Short-Run Model

The model describes a global economy consisting of two regions (labeled $j \in \{1,2\}$) and three production sectors. One produces an intermediate good 'transport service' ξ for the industrial sectors by employing a fixed amount of immobile unskilled work force L. A second sector is manufacturing, denoted by the symbol M, which produces a continuum of i varieties of a horizontally-differentiated final good through mobile human capital H and transport ξ as input factors. A third sector is an aggregated sector, denoted by the symbol Q, which produces a homogeneous traditional final good using only immobile unskilled labor L. M is characterized by increasing returns and monopolistic competition *à la* Dixit and Stiglitz (1977). Because of consumer preferences for variety and increasing returns to scale, each firm specializes in producing a distinct variety of the manufactured good. Hence, the total number of active firms in the two-region economy, $N = n_1 + n_2$ equals the number of varieties available in the market. The traditional and the transportation sectors produce under Walrasian conditions (constant returns to scale and perfect competition). The traditional good is chosen as the *numéraire* (i.e. its price is set at unity).

For the purpose of assessing environmental sustainability as related to the use of space, we explicitly model a pollutive transportation sector as a variable input of production. This comes down to assuming that, unlike the literature following Krugman (1991), the domestic trade of the manufacturing good is costly. In line with a Krugman-like modeling setting, international trade of the composite manufacturing good occurs at a certain cost, whereas trade costs are zero for both inter- and intra-regional shipment of the traditional good. $L = L_1 + L_2$ and $H = H_1 + H_2$ denote the total of unskilled and skilled laborers, respectively. In the initial spatial setting, skilled workers are unevenly distributed across the two regions; the share of skilled workers living in region 1 is denoted by *h*, with $h = H_1/H$. Unskilled workers, on the other hand, are assumed to be evenly spread across regions, so that $L_j = L/2$. Each unskilled worker supplies one unit of labor.

1.1 Households

Workers maximize utility by consuming the two goods and suffer from negative effects on utility because of external environmental effects associated with economic activity. Aggregate utility is a Cobb-Douglas function of consumption of the traditional commodity *Q* and consumption of the aggregate manufactured good *M*. The latter is modeled as a CES function of consumption levels $c_{jj}(i)$ and $c_{kj}(i)$ of a particular variety *i* of the manufactured good that is sold in region *j* and produced in regions *j* and *k*.[3]

The negative effect of the environmental externality on utility is captured by a multiplicative term $\Theta\left(E_j^L\right)$. Many earlier studies employed an additive functional form to achieve analytical results [e.g. Rauscher (2003); Lange and Quaas (2007); Elbers and Withagen (2004)]. This comes down to assuming constant marginal disutility associated with the environmental externality. Unlike these studies we treat the environmental externality as part of a multiplicative utility function, which ensures a more realistic relationship between pollution and utility, while still allowing for analytical solutions of the model. This modeling choice is moreover in line with a theoretical study of appropriate functional forms to describe environmental externalities [Ebert and Welsch (2004)].

$$U_j = M_j^{\delta} Q_j^{1-\delta} \Theta\left(E_j^L\right),\ j = (1,2); \Theta\left(E_j^L\right) \le 1, \text{ with:}$$

$$M_j = \left[\int_{i=0}^{n_j} c_{jj}(i)^{(\varepsilon-1)/\varepsilon} di + \int_{i=0}^{n_k} c_{kj}(i)^{(\varepsilon-1)/\varepsilon} di\right]^{\frac{\varepsilon}{\varepsilon-1}}; j,k = 1,2; j \ne k; i \in N. \tag{1}$$

[3] For ease of notation, we drop the index *i* for varieties in the remainder of the paper.

Here, $0<\delta<1$ is the share of income Υ_j spent on manufactures; $\varepsilon>1$ is the elasticity of substitution between varieties; and $\Theta\left(E_j^L\right)$ is the damage function associated with local flows of pollution E_j^L which alters individuals' utility in *j*.

Domestic consumption of traded goods c_{kj} results from standard utility maximization:

$$c_{kj}=\frac{(p_{kj})^{-\varepsilon}}{I_j^{1-\varepsilon}}\delta\Upsilon_j, j,k=1,2; j\neq k. \tag{2}$$

Here p_{kj} is the price of a good produced in *k* and consumed in *j*, and $I_j=[n_j p_j^{1-\varepsilon}+n_k p_{kj}^{1-\varepsilon}]^{1/(1-\varepsilon)}$ is Dixit-Stiglitz's (1977) price index of the manufactured good in *j*.

1.2 Firms

Manufacturing firms produce using both skilled labor *H* and domestic transport *ξ* as inputs. Skilled workers are hired at a domestic wage rate w_j, while domestic transport services are paid a price p^{ξ} independent of the region *j* considered. The cost structure of a typical *j*-firm which produces a quantity x_j of the manufactured good entails fixed costs in human capital, αw_j, and variable costs in terms of transport requirements per unit of output, $p^{\xi}\xi_j$:

$$\chi_j=\alpha w_j+p^{\xi}\xi_j x_j. \tag{3}$$

Trade also occurs between the two regions. To avoid modeling a separate interregional transportation sector, we use the 'iceberg' form of transport costs associated with the interregional trade of manufactured goods (Samuelson, 1952). This means that if a variety of the manufactured good produced in location *j* is sold in the same region at price p_{jj} then it will be charged a price p_{jk} in consumption location *k* that satisfies $p_{jk}=p_{jj}T_{jk}$. Here $T_{jk}>1$ is the iceberg unitary trade cost of the manufactured good, which represents the number of goods sent per unit received. We assume that interregional trade costs are the same in each direction, $T=T_{jk}=T_{kj}$.

Next we formalize an agglomeration effect and consider the impact of spatial clustering of economic activities on the transport intensity of production in the manufacturing sector. This effect occurs through a two mechanisms. The first one is based on the regional endowment of infrastructure, which represents the spatial form (or structure) of the market;

and the second mechanism is related to the number of firms in the spatial form, which determines the density of economic activity. The region specific transport requirement results from the *endogenous* agglomeration effect, as captured by variable ξ_j:

$$\xi_j = \beta_j \bar{\psi}\left(n_j\right). \tag{4}$$

Here the parameter $\beta_j > 0$ captures the impact of regional spatial form, related to and captured by the regional infrastructure endowment, on the transport intensity of regional production in manufacturing. In addition, $0 \le \bar{\psi}\left(n_j\right) \le 1$ is the equivalent impact of market density, which is a function of the number of firms that are active in the regional market when its spatial extension is determined. In the remainder of the paper, we refer to β_j as the 'market-form' effect and to $\bar{\psi}\left(n_j\right)$ as the 'market-density' effect.

Parameter β_j represents the transport requirement per unit of output in production. We can think of this as capturing (being inversely related to) the degree of 'urbanization' of a given spatial economy, or the spatial concentration of domestic transportation and telecommunications infrastructure networks. Since infrastructure is characterized by slow dynamics or inertia, β_j is treated as an exogenous parameter. Two possible spatial forms (or structures) for each region are considered: namely, a spatially-developed organization of manufacturing activities, with a high intensity of infrastructure development (urbanized space), and a less intense use of space by these activities on (undeveloped) land. We consider a two-region system, which then gives rise to three possible spatial configurations of the global economy (urban + undeveloped; urban + urban; and undeveloped + undeveloped).[4]

The multiplicative term $\bar{\psi}\left(n_j\right)$ captures the impact of the market density on intra-industry transaction (communication and transport) costs and the technological spillovers.[5] In line with empirical evidence on the effect of density of the economic activity on the structure of production (Ciccone and Hall, 1996; Keller, 2002; Duranton and Puga, 2004), we posit $\bar{\psi}\left(0\right) = 1$ to indicate no positive effect of agglomeration on production costs in the absence of

[4] Actually, with the two possible regional structures described, $2^2 = 4$ spatial configurations for the two-region economy are possible. However, two of these are spatial mirror images of each other.

[5] The 'market-density' external effect that we model acts so as to reduce the average cost of production at the industry level, thus overriding the firm scale. As such it can be identified with external economies in the sense of Scitovsky (1954).

firms; and $\overline{\psi}'\left(n_j\right) < 0$ to mean that the higher the number of firms, the stronger the reduction in production costs.

Given eq. (3), profit-maximization leads to mark-up pricing for the manufactured good:

$$p_j = \frac{\varepsilon}{\varepsilon - 1} p^{\xi} \xi_j. \tag{5}$$

The traditional good and the transport service commodity are produced using unskilled labor as a linear input.[6]

Production in the traditional sector is assumed to have a one-to-one relationship with unskilled labor and final product, whereas in transport service supply the labor requirement per unit of output is captured by parameter γ. We take the wage of unskilled workers as the *numéraire*.[7] Marginal cost pricing in the domestic transportation sector then implies:

$$p^{\xi} = \gamma. \tag{6}$$

The domestic supply of the traditional good is:

$$Q_j = L/2 - \gamma \xi_j n_j x_j, \tag{7}$$

where the second term on the right-hand side of (7) represents the effect of unskilled workers being employed in the transport sector [see eq.(3)].

1.3 Market Equilibrium

For a given regional distribution of the skilled labor factor H_j, the short-run model is determined by a set of four equations (for details, see Grazi et al, 2007).

$$\Upsilon_j = w_j H_j + L/2. \tag{8}$$

Here, Υ_j is the income generated in each region by w_j, the wage rate of skilled workers, H_j, and the *numéraire* wage of L_j unskilled workers.

[6] The assumption of linearity in the traditional/agricultural constant returns sector is very standard (Krugman, 1991). We extend it to the transport service sector in order to keep the analysis simple.

[7] This is a consequence of assuming free trade for the *numéraire* traditional good *Q*, which in turn comes down to its price being equal to 1 across regions: $p_j^Q = p_k^Q = p^Q = 1$, with $j,k = \{1,2\}$. Marginal cost pricing implies the interregional equalization of the wages of unskilled labor input *L* used in the traditional sector: $p^Q = w^L = 1$.

$$n_j = \frac{H_j}{\alpha}, \tag{9}$$

where a fixed input requirement α indicates that the total number of firms operating in region j, n_j, is proportional to locally available skilled laborers.

As a consequence of the profit maximization behavior in a monopolistically competitive market, in both regions firms will enter and exit the manufacturing sector until the point at which profits are zero. Therefore, by substituting (5) into the profit function $\pi_j = p_j x_j - \chi_j$ and setting $\pi_j = 0$, the wage rate w_j at the equilibrium is:

$$w_j = \frac{\gamma \xi_j x_j}{\alpha(\varepsilon - 1)}. \tag{10}$$

The market-clearing size of a typical firm in equilibrium is $x_j = c_{jj} + Tc_{jk}$. Substituting (2), (5) and (6) in (10) gives equilibrium solutions for x_j and I_j:

$$x_j = \delta \left[\frac{\varepsilon - 1}{\gamma \xi_j \varepsilon} \right] \left(\frac{\Upsilon_j}{I_j^{1-\varepsilon}} + \frac{\phi \Upsilon_k}{I_k^{1-\varepsilon}} \right), \text{ with}$$

$$I_j = \frac{\varepsilon}{\varepsilon - 1} \gamma \left(n_j \xi_j^{1-\varepsilon} + \phi n_k \xi_k^{1-\varepsilon} \right), 0 \leq \phi \leq 1. \tag{11}$$

Here $\phi = T^{1-\varepsilon}$ is a measure of the openness to interregional trade, with $\phi = 0$ representing maximal barriers to interregional trade (or autarky), and $\phi = 1$ free trade across regions.[8]

Given eq. (9) and recalling that the share of the regional population equals $h = H_1 / H$, the 'market density' effect $\bar{\psi}(n_j)$ in (4) can be re-written as a function of *h*: $\psi(h)$ in region 1 and $\psi(1-h)$ in region 2. Moreover, substituting equations (8), (9) and (11) into (10), and using the definition of the regional share of population *h*, the model can be analytically solved in the regional wage levels w_1 and w_2:

[8] In the NEG approach, transport costs allow one to study the extent to which space affects economic decisions by individual agents (consumers and producers), and how these decisions in turn drive the spatial distribution of economic activities.

$$w_1 = \frac{\delta/\varepsilon}{1-\delta/\varepsilon}\frac{L}{2}\beta_1^{1-\varepsilon}\psi(h)^{1-\varepsilon}\cdot$$
$$\cdot\frac{2\phi\beta_1^{1-\varepsilon}\psi(h)^{1-\varepsilon}h + \left[1-\frac{\delta}{\varepsilon}+\left(1+\frac{\delta}{\varepsilon}\right)\phi^2\right]\beta_2^{1-\varepsilon}\psi(1-h)^{1-\varepsilon}\left(1-h\right)}{\phi\left[\beta_1^{2(1-\varepsilon)}\psi(h)^{2(1-\varepsilon)}h^2+\beta_2^{2(1-\varepsilon)}\psi(1-h)^{2(1-\varepsilon)}(1-h)^2\right]+\beta_1^{1-\varepsilon}\beta_2^{1-\varepsilon}\psi(h)^{1-\varepsilon}\psi(1-h)^{1-\varepsilon}h(1-h)};$$

(12)

$$w_2 = \frac{\delta/\varepsilon}{1-\delta/\varepsilon}\frac{L}{2}\beta_2^{1-\varepsilon}\psi(1-h)^{1-\varepsilon}\cdot$$
$$\cdot\frac{2\phi\beta_2^{1-\varepsilon}\psi(1-h)^{1-\varepsilon}\left(1-h\right) + \left[1-\frac{\delta}{\varepsilon}+\left(1+\frac{\delta}{\varepsilon}\right)\phi^2\right]\beta_1^{1-\varepsilon}\psi(h)^{1-\varepsilon}h}{\phi\left[\beta_1^{2(1-\varepsilon)}\psi(h)^{2(1-\varepsilon)}h^2+\beta_2^{2(1-\varepsilon)}\psi(1-h)^{2(1-\varepsilon)}(1-h)^2\right]+\beta_1^{1-\varepsilon}\beta_2^{1-\varepsilon}\psi(h)^{1-\varepsilon}\psi(1-h)^{1-\varepsilon}h(1-h)}.\quad \text{(12bis)}$$

1.4 Local Pollution Externalities

The small literature that exists on NEG with agglomeration and environmental externalities considers the local effect of pollution (flow), meaning the (immediate) negative impact on the utility of individuals living in the respective region [Rauscher (2003); van Marrewijk (2005); Lange and Quaas (2007)]. Like these studies, we initially assume that the environmental externality (pollution) is local and only generated by manufacturing.[9] Yet, unlike these studies, we reject the standard assumption of proportionality of pollution to the output x_j and instead consider pollution as a function of transport-related energy use by the *j*-manufacturing sector $\xi_j x_j$. By so doing, we are able to address the relationship between energy use, energy intensity and production structure.

Local pollution emissions by manufacturing production in region *j* which affect utility in *j* are then generated in the following manner:

$$E_j^L = a^L\left(\xi_j n_j x_j\right). \tag{13}$$

[9] Later, in Section IV, we relax this assumption to consider global environmental externalities that arise from interregional trade/transport as well, and which is functional to the aim of addressing sustainability. Since global environmental externalities do not introduce heterogeneity across regions, they have no influence on the dynamics of migration. Hence, they can be omitted in this first-step analysis, which aims to determine the patterns and stability of the spatial long-run equilibria.

Here, a^L represents the intensity of externalities generated by the transport input ξ_j.[10] The pollution term in utility, $\Theta\left(E_j^L\right)$ [see eq. (1)] captures the effect of pollution externalities on utility. We posit $\Theta\left(0\right)=1$ to indicate no negative effect of pollution on utility in the absence of any flow of pollution; and $\Theta'\left(E_j^L\right)<0$ to mean that the higher the pollution level, the stronger is its negative effect on utility.

2. The Long-Run Model and the Dynamics of Migration

Next, we study the long-term impact of different spatial configurations on production allocation when agglomeration- and local pollution-related effects matter. As in the "footloose entrepreneur" framework [Forslid and Ottaviano (2003)] the model dynamics is driven by international migration of individuals belonging to the skilled population. The resulting spatial equilibria are defined over the share h of skilled workers living in region 1, where $h=H_1/H$. Then the study of dynamic behavior of the core model variables is carried out for different values of trade barrier, ϕ. Consequently, all the variables in the dynamics analysis (wage, price index, pollution externality, etc.), can be expressed as functions of variables h and ϕ ($w_j(h,\phi)$, $I_j(h,\phi)$, $E_j^L(h,\phi)$, etc.).

The dynamics of migration and resulting spatial equilibria follow from individuals comparing wages at different locations, as captured by the indirect utility differential between region 1 and 2 as $\Omega(h,\phi)=V_1(h,\phi)-V_2(h,\phi)$, where the indirect utility V_j associated with (1) is specified as:

$$V_j(h,\phi)=\Gamma\frac{w_j(h,\phi)}{I_j(h,\phi)^{\delta}}\Theta(h,\phi),\ j=\{1,2\}. \tag{14}$$

Here, $\Gamma=\delta^{\delta}(1-\delta)^{1-\delta}$ is a constant that depends on the share of income devoted to manufacturing good purchases, δ.

Substituting (14) in the indirect utility differential $\Omega(h,\phi)$ gives the following derived relationship, which represents the incentive to move from region 2 to region 1:

[10] An alternative way to model local pollution would be to relate this to the marginal input factor unskilled labor, as in Copeland and Taylor (2004). Our formulation, although simple, has the advantage of specifying explicitly the impact mechanism via the transport-related energy use, and addressing the impact of agglomeration on emission intensity.

$$\Omega(h,\phi)=\Gamma\left[\frac{w_1(h,\phi)}{I_1(h,\phi)^{\delta}}\Theta(h,\phi)-\frac{w_2(h,\phi)}{I_2(h,\phi)^{\delta}}\Theta(1-h,\phi)\right]. \tag{15}$$

Given $h\in[0;1]$, the equation describing the dynamics of factor mobility can be expressed as follows:[11]

$$\frac{dh}{dt}=\begin{cases}\Omega(h,\phi), & \text{if } 0<h<1\\ \max(0,\Omega(h,\phi)), & \text{if } h=0\\ \min(0,\Omega(h,\phi)), & \text{if } h=1\end{cases}. \tag{16}$$

Clearly, a long-run spatial equilibrium is defined by condition:

$$\frac{dh}{dt}=0. \tag{17}$$

Substituting (15) and (16) into (17) gives the implicit relationship between the distribution of population h and the trade barrier ϕ in the long run. Such an equilibrium is stable only if $\frac{\partial\Omega}{\partial h}(h,\phi)<0$. For a given spatial configuration, a certain pattern of population distribution associated with a trade barrier level ϕ defines a stable long-run equilibrium if one of the three following conditions holds:

$$a)\begin{cases}0<h<1\\ \Omega(h,\phi)=0,\dfrac{\partial\Omega}{\partial h}(h,\phi)<0\end{cases};\ b)\begin{cases}h=1\\ \Omega(h,\phi)\geq 0\end{cases};\ c)\begin{cases}h=0\\ \Omega(h,\phi)\leq 0\end{cases}. \tag{18}$$

II. Equilibrium with Agglomeration and Pollution Effects

This section analytically investigates the stability of long-run spatial equilibria once agglomeration and pollution flow effects are accounted for, as in condition (18). By

[11] Note that dynamics are implicit-in-time in this type of modeling framework [Krugman (1991)]. This allows us to omit the index for time dependence from the variables of the long-run model.

substituting (12), (12*bis*), and (11) into (15), the latter can be rewritten as:

$$\Omega(h,\phi)=\frac{\dfrac{\Gamma'}{\beta_1^{\delta}}\,\omega(h,\phi)}{\phi\left[h^2\psi(h)^{2(1-\varepsilon)}+(1-h)^2\left(\psi(1-h)\dfrac{\beta_2}{\beta_1}\right)^{2(1-\varepsilon)}\right]+\left[1-\dfrac{\delta}{\varepsilon}+\phi^2\left(1+\dfrac{\delta}{\varepsilon}\right)\right]h(1-h)\left(\dfrac{\beta_2}{\beta_1}\psi(h)\psi(1-h)\right)^{1-\varepsilon}}. \tag{19}$$

Here $\Gamma'=\Gamma\dfrac{\delta/\varepsilon}{1-\delta/\varepsilon}\dfrac{L}{2}\left[\dfrac{(\varepsilon-1)\alpha^{\frac{1}{1-\varepsilon}}}{\gamma\varepsilon}\right]^{\delta}$ is a positive parameter and ω is a function that depends on variables ϕ and h in the following way:

$$\omega(h,\phi)=\frac{\psi(h)^{1-\varepsilon}\left\{2h\phi\psi(h)^{1-\varepsilon}+\left(\dfrac{\beta_2}{\beta_1}\right)^{1-\varepsilon}\left[1-\dfrac{\delta}{\varepsilon}+\phi^2\left(1+\dfrac{\delta}{\varepsilon}\right)\right](1-h)\psi(1-h)^{1-\varepsilon}\right\}}{\left[h\psi(h)^{1-\varepsilon}+\left(\dfrac{\beta_2}{\beta_1}\right)^{1-\varepsilon}\phi(1-h)\psi(1-h)^{1-\varepsilon}\right]^{\frac{\delta}{1-\varepsilon}}}\Theta(h,\phi)+$$

$$-\frac{\left(\dfrac{\beta_2}{\beta_1}\right)^{1-\varepsilon}\psi(1-h)^{1-\varepsilon}\left\{2\phi\left(\dfrac{\beta_2}{\beta_1}\right)^{1-\varepsilon}(1-h)\psi(1-h)^{1-\varepsilon}+\left[1-\dfrac{\delta}{\varepsilon}+\phi^2\left(1+\dfrac{\delta}{\varepsilon}\right)\right]h\psi(h)^{1-\varepsilon}\right\}}{\left[h\phi\psi(h)^{1-\varepsilon}+\left(\dfrac{\beta_2}{\beta_1}\right)^{1-\varepsilon}(1-h)\psi(1-h)^{1-\varepsilon}\right]^{\frac{\delta}{1-\varepsilon}}}\Theta(1-h,\phi). \tag{20}$$

Note that the formulation in (19) generalizes the result obtained by Forslid and Ottaviano (2003) in three ways: *i*) it allows for a variable positive spatial spillover effect associated with the size of market (i.e. $\psi(h)\leq 1$); *ii*) it includes an environmental externality that negatively affects the utility of individuals (i.e. $\Theta(h,\phi)\leq 1$); and *iii*) it enables to represent *ex-ante* differences in the regional spatial setting (i.e. β_1 and β_2 may take different values). On the other hand, setting $\psi(h)=1$, $\Theta(h,\phi)=1$ and $\beta_1=\beta_2$ in equation (19) produces the same results as Forslid and Ottaviano (2003).

To be maximally consistent with the original framework by Forslid and Ottaviano (2003), we start this section's analysis by considering the case of *ex-ante* identical regions ($\beta_1=\beta_2$), and derive for each the conditions that drive the nature of the long-run equilibria (sub-section II.1). The study of equilibria is then extended to consider different spatial settings for the regions ($\beta_1\neq\beta_2$) (sub-section II.2). Finally, the results are analyzed and discussed in the light of the interplay between agglomeration and environmental drivers (sub-section II.3).

Adapting the expressions in (19) and (20) to the range of possible spatial settings of the economy (as defined by the combinations $[\beta_1;\beta_2]$), we successively investigate the stability conditions of the core-periphery ($h=1$) and symmetric spreading ($h=0.5$) equilibria, as is standard in the NEG literature. The key innovation that moves this section beyond the previous NEG literature is that it analytically derives the general stability conditions of the partial agglomeration equilibria ($0.5<h<1$).[12]

1. Equilibrium of Symmetric Spatial Configurations

Here we study the long-run equilibria associated with (18) in symmetric spatial configurations of the two-region economy (i.e. $\beta_j=\beta_k=\beta$). We limit the analysis to the case $0.5\le h\le 1$, since the findings are symmetrical around $h=0.5$. For these conditions, the indirect utility in (19) and the function in (20) can be rewritten as follows:

$$\Omega(h,\phi)=\frac{\Gamma'}{\beta^{\delta}}\frac{\omega(h,\phi)}{\phi\left[h^2\psi(h)^{2(1-\varepsilon)}+(1-h)^2\psi(1-h)^{2(1-\varepsilon)}\right]+\left[1-\frac{\delta}{\varepsilon}+\phi^2\left(1+\frac{\delta}{\varepsilon}\right)\right]h(1-h)\psi(h)^{1-\varepsilon}\psi(1-h)^{1-\varepsilon}}, \tag{21}$$

and

$$\begin{aligned}\omega(h,\phi)=&\frac{\psi(h)^{1-\varepsilon}\left\{2h\phi\psi(h)^{1-\varepsilon}+\left[1-\frac{\delta}{\varepsilon}+\phi^2\left(1+\frac{\delta}{\varepsilon}\right)\right](1-h)\psi(1-h)^{1-\varepsilon}\right\}}{\left[h\psi(h)^{1-\varepsilon}+\phi(1-h)\psi(1-h)^{1-\varepsilon}\right]^{\frac{\delta}{1-\varepsilon}}}\Theta(h,\phi)+\\&-\frac{\psi(1-h)^{1-\varepsilon}\left\{2\phi(1-h)\psi(1-h)^{1-\varepsilon}+\left[1-\frac{\delta}{\varepsilon}+\phi^2\left(1+\frac{\delta}{\varepsilon}\right)\right]h\psi(h)^{1-\varepsilon}\right\}}{\left[h\phi\psi(h)^{1-\varepsilon}+(1-h)\psi(1-h)^{1-\varepsilon}\right]^{\frac{\delta}{1-\varepsilon}}}\Theta(1-h,\phi).\end{aligned} \tag{22}$$

Next we derive the analytical equilibrium conditions:

1.1 Core-Periphery Pattern in Symmetric Configurations

Rewriting equation (20) for the case of a Core-Periphery (CP) pattern ($h=1$) gives:[13]

[12] Lange and Quaas (2007) also obtain partial equilibria, but only for a restricted set of the trade parameter values. Such a limitation excludes any realistic application of their model findings to environmental policy analyses.

[13] Deriving $E_j^L(1)$ from (10), (12), (12*bis*) and (13) shows that the pollution flow function is independent from ϕ for $h=1$. Hence the index ϕ in $\Theta(1)$ can be omitted.

$$\omega(1,\phi) = \frac{\psi(1)^{1-\varepsilon}}{\left[\psi(1)^{1-\varepsilon}\right]^{\frac{\delta}{1-\varepsilon}}}\left\{2\phi\Theta(1)\psi(1)^{1-\varepsilon} - \frac{\left[1-\frac{\delta}{\varepsilon}+\phi^2\left(1+\frac{\delta}{\varepsilon}\right)\right]\Theta(0)\psi(0)^{1-\varepsilon}}{\phi^{\frac{\delta}{1-\varepsilon}}}\right\}.$$

By assumption, because regional economic activity is absent, $\Theta(0)=1$ and $\psi(0)=1$, that is, these variables take their maximum values. Combining this with the condition $\omega(1,\phi)>0$ (18-*b*), which assures that the core-periphery pattern ($h=1$) is a stable equilibrium for any ϕ value, we obtain $\Theta(1)\psi(1)^{1-\varepsilon} \geq \sigma_{\mathrm{CP}}(\phi)$, where:

$$\sigma_{\mathrm{CP}}(\phi) = \frac{1-\frac{\delta}{\varepsilon}+\phi^2\left(1+\frac{\delta}{\varepsilon}\right)}{2(\phi)^{1+\frac{\delta}{1-\varepsilon}}}. \tag{23}$$

This a concave function that defines the stability condition of the core-periphery equilibrium pattern in the case of symmetric configurations, as summarized by the following condition:

CONDITION 1: Given $\sigma_{\mathrm{CP}}(\phi)$ in (23), the core-periphery pattern ($h=1$) is a stable equilibrium for a trade barrier ϕ ($0 \leq \phi \leq 1$) if and only if: $\Theta(1)\psi(1)^{1-\varepsilon} \geq \sigma_{\mathrm{CP}}(\phi)$.[14]

The possible stable equilibrium outcomes are summarized by the following proposition:

PROPOSITION 1: Given $\Theta(1)$, $\psi(1)$, and $\sigma_{\mathrm{CP}}^{\min} = \frac{\varepsilon-1}{\varepsilon}\left(\frac{\varepsilon-\delta}{\varepsilon-1+\delta}\right)^{\frac{1}{2}\left(\frac{\delta}{\varepsilon-1}+1\right)}\left(\frac{\varepsilon-1-\delta}{\varepsilon+\delta}\right)^{\frac{1}{2}\left(\frac{\delta}{\varepsilon-1}-1\right)}$, three cases must be distinguished according to the position of $\Theta(1)\psi(1)^{1-\varepsilon}$ with respect to $\sigma_{\mathrm{CP}}^{\min}$ and 1:

CP–*i*: If $\Theta(1)\psi(1)^{1-\varepsilon} \leq \sigma_{\mathrm{CP}}^{\min}$, the full agglomeration is never an equilibrium, whatever the trade freeness;

CP–*ii*: If $\sigma_{\mathrm{CP}}^{\min} < \Theta(1)\psi(1)^{1-\varepsilon} < 1$, the full agglomeration is a stable equilibrium for intermediate trade freeness $\phi \in \left[\underline{\phi}_S;\overline{\phi}_S\right]$, while it is unstable for $\phi \in [0;1] \setminus \left[\underline{\phi}_S;\overline{\phi}_S\right]$;

[14] Stability of long-run equilibrium in a NEG model requires that the additional so-called "no black hole" condition is satisfied, which imposes that the full agglomeration is never a stable equilibrium in case of autarky $\phi=0$ [Fujita, Krugman and Venables (1999)]. According to Condition 1, this means ensuring that function $\sigma_{\mathrm{CP}}(\phi)$ tends to infinity when $\phi=0$, which, from (23), is in turn equivalent to $1+\delta/(1-\varepsilon)>0$. This analytical form for the "no black hole" condition is similar to the one obtained in Forslid and Ottaviano (2003), and is supposed to hold in the reminder of the paper.

CP–*iii*: If $\Theta(1)\psi(1)^{1-\varepsilon} \geq 1$, the full agglomeration is a stable equilibrium for a sufficiently high value of trade freeness $\phi \in [\phi_S;1]$, while it is unstable for $\phi \in [0;\phi_S]$;

The threshold point ϕ_S (with $\underline{\phi}_S$ and $\overline{\phi}_S$ indicating its upper and lower value in case of existence of multiple points) is the "sustain point" in the sense of Fujita, Krugman and Venables (1999). It is implicitly given by any ϕ value that satisfies condition $\Theta(1)\psi(1)^{1-\varepsilon} = \sigma_{\text{CP}}(\phi)$.

(See Appendix B.1 for a proof).

We can compare our results with those obtained by Lange and Quaas (2007). Even though their model differs from ours in the specification of the negative externalities in the utility function (additive vs. multiplicative), the results of Proposition 1 are comparable with theirs in the absence of the endogenous agglomeration-driving 'market-density' effect. In our analysis this translates into $\psi(1) = 1$. Only two out of the above three possible outcomes then emerge:[15] *i*) if $\Theta(1) \leq \sigma_{\text{CP}}^{\min}$, the core-periphery structure is unstable independently of freeness of trade (case CP–*i*); and *ii*) if $\Theta(1) > \sigma_{\text{CP}}^{\min}$, two "sustain points" $\underline{\phi}_S$ and $\overline{\phi}_S$ exist and the full agglomeration is a stable equilibrium only for intermediate trade freeness $\phi \in \left[\underline{\phi}_S;\overline{\phi}_S\right]$ (case CP–*ii*).

1.2 Symmetric-Spreading Pattern in Symmetric Configurations

Let us now turn to consider the stability range of the core-periphery equilibrium $h = 0.5$. Rewriting (20) for this specific case shows that $\omega(0.5,\phi) = 0$ for any ϕ, so that the symmetric outcome is always an equilibrium. Such an equilibrium is stable if and only if $\frac{\partial\Omega}{\partial h}(0.5,\phi) \leq 0$.

We introduce $d_{\psi}^{(0.5)} = -2\frac{\psi'(0.5)}{\psi(0.5)}$ and $d_{\Theta}^{(0.5)} = -2\frac{\frac{\partial\Theta}{\partial E}(0.5)}{\Theta(0.5)}$ as a measure of the intensity of the agglomeration and the environmental effects at $h = 0.5$, respectively.[16] Using (21) and (22) we can derive:

---

[15] Note that case CP–*iii* in Proposition 1 never emerges in the absence of the market density effect because condition $\Theta(1) > 1$ then does not hold.

[16] Deriving $E_j^L(0.5)$ from (10), (12), (12*bis*) and (13) shows that the pollution flow function is independent of ϕ for $h = 0.5$. Hence, the index ϕ in $\Theta(0.5)$ and in $d_{\Theta}^{(0.5)}$ can be omitted. Moreover, since ψ and Θ are both decreasing in *h*, their derivatives are negative. Consequently, $d_{\psi}^{(0.5)}$ and $d_{\Theta}^{(0.5)}$ are positive terms.

$$\frac{\partial\Omega}{\partial h}\left(0.5,\phi\right)=\frac{\left[\frac{1}{\alpha}\left(\frac{\varepsilon\gamma}{\varepsilon-1}\right)^{1-\delta}\right]\Theta(0.5,\phi)2^{1-\frac{\delta}{\varepsilon-1}}L\delta\left[\beta^{1-\varepsilon}(1+\phi)\psi(0.5)^{\frac{\delta}{\varepsilon-1}}\right]}{\left(\varepsilon-\delta\right)\left(\phi+1\right)\left[\delta\left(\phi-1\right)+\varepsilon\left(\phi+1\right)\right]}\sigma_{\mathrm{SS}}(\phi), \tag{24}$$

where $\sigma_{\mathrm{SS}}(\phi)$ can be decomposed into three components:

$$\sigma_{\mathrm{SS}}(\phi)={}^{(\mathrm{FE})}\,\sigma_{\mathrm{SS}}(\phi)+d_{\psi}^{(0.5)}\sigma_{\mathrm{SS}}^{(\psi)}(\phi)-d_{\Theta}^{(0.5)}\sigma_{\mathrm{SS}}^{(\Theta)}(\phi), \tag{25}$$

with

$$\begin{aligned}
{}^{(\mathrm{FE})}\sigma_{\mathrm{SS}}(\phi)&=2\left(1-\phi\right)\left[\varepsilon+\delta+\frac{\delta}{\varepsilon-1}\left(\delta+\varepsilon\right)\right]\left[\phi-\frac{\varepsilon-\delta}{\varepsilon+\delta}\frac{\varepsilon-1-\delta}{\varepsilon-1+\delta}\right];\\
\sigma_{\mathrm{SS}}^{(\psi)}(\phi)&=\frac{1}{2}\left\{-\delta\phi^{2}\left(\varepsilon+\delta\right)+\phi\left[4\varepsilon\left(\varepsilon-1\right)+2\delta^{2}\right]+\delta\left(\varepsilon-\delta\right)\right\};\\
\sigma_{\mathrm{SS}}^{(\Theta)}(\phi)&=\frac{\delta\varepsilon a^{f}L\left(\varepsilon-1\right)}{\gamma\left(\varepsilon-\delta\right)}\phi\left(2+d_{\psi}^{(0.5)}\frac{\varepsilon-1}{2}\right).
\end{aligned} \tag{26}$$

Since all other terms on the right-hand side of equation (24) are positive, $\frac{\partial\Omega}{\partial h}\left(0.5,\phi\right)$ has the same sign as $\sigma_{\mathrm{SS}}(\phi)$. The stability condition of the symmetric-spreading equilibrium in symmetric configurations can then be expressed as follows:

CONDITION 2: A symmetric distribution of skilled workers ($h=0.5$) is always an equilibrium. Given (24), such an equilibrium is stable if and only if $\sigma_{\mathrm{SS}}(\phi)<0$.

The conditions for determining the sign of $\sigma_{\mathrm{SS}}(\phi)$ are analytically obtained from (25) and (26). Given Condition 2, this allows to explicitly investigate the stability conditions of the symmetric spreading equilibrium, as summarized in the following proposition:

PROPOSITION 2: Let the following functions be defined:

$$\begin{aligned}
d_{\psi,0}&=\frac{4\left(\varepsilon-1-\delta\right)}{\delta(\varepsilon-1)};\\
\zeta\left(d_{\psi}^{(0.5)}\right)&=\frac{4\gamma\left(\varepsilon-\delta\right)}{\delta a^{f}L}\frac{\left(\varepsilon-\delta\right)}{4+d_{\psi}^{(0.5)}\left(\varepsilon-1\right)};\\
\zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)&=\zeta\left(d_{\psi}^{(0.5)}\right)+\frac{2\gamma\left(\varepsilon-\delta\right)\left\{4\left(\varepsilon-1\right)+\varepsilon\left[4+d_{\psi}^{(0.5)}\left(\varepsilon-1\right)\right]\right\}^{2}}{\varepsilon\left(\varepsilon-1\right)^{2}\delta a^{f}L\left[4+d_{\psi}^{(0.5)}\left(\varepsilon-1\right)\right]}\cdot
\end{aligned}$$

$$\cdot\frac{\delta^{2}}{4\delta^{2}+4\varepsilon\left(\varepsilon-1\right)+\delta^{2}d_{\psi}^{(0.5)}\left(\varepsilon-1\right)+\sqrt{\left(\varepsilon^{2}-\delta^{2}\right)\left\{\left[4\left(\varepsilon-1\right)\right]^{2}-\delta^{2}\left[4+d_{\psi}^{(0.5)}\left(\varepsilon-1\right)\right]^{2}\right\}}}.$$

Then stability of the symmetric-spreading equilibrium depends on these functions, where five cases can be distinguished:

SS–*i*: If $d_{\psi}^{(0.5)} > d_{\psi,0}$ and $d_{\Theta}^{(0.5)} < \zeta\left(d_{\psi}^{(0.5)}\right)$, the symmetric-spreading is never a stable equilibrium;

SS–*ii*: If $d_{\psi}^{(0.5)} < d_{\psi,0}$ and $d_{\Theta}^{(0.5)} < \zeta\left(d_{\psi}^{(0.5)}\right)$, a value ϕ_b exists so that the symmetric-spreading is a stable equilibrium for all $\phi \in [0;\phi_b]$;

SS–*iii*: If $d_{\psi}^{(0.5)} > d_{\psi,0}$ and $d_{\Theta}^{(0.5)} > \zeta\left(d_{\psi}^{(0.5)}\right)$, a threshold value ϕ_b exists so that the symmetric spreading is a stable equilibrium for all $\phi \in [\phi_b;1]$;

SS–*iv*: If $d_{\psi}^{(0.5)} < d_{\psi,0}$ and $\zeta\left(d_{\psi}^{(0.5)}\right) < d_{\Theta}^{(0.5)} < \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, two threshold values exist, $\underline{\phi}_b$ and $\overline{\phi}_b$, such that the symmetric equilibrium is stable for all $\phi \in \left[0;\underline{\phi}_b\right] \cup \left[\overline{\phi}_b;1\right]$ and unstable for all $\phi \in \left[\underline{\phi}_b;\overline{\phi}_b\right]$

SS–*v*: If $d_{\psi}^{(0.5)} < d_{\psi,0}$ and $d_{\Theta}^{(0.5)} > \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, the symmetric-spreading equilibrium is always stable.

The threshold point ϕ_b (whose lower and upper values are represented by $\underline{\phi}_b$ and $\overline{\phi}_b$, respectively) appearing in cases SS–*ii*, SS–*iii* and SS–*iv* is the "break point" in the sense of Fujita, Krugman and Venables (1999). It is implicitly given by condition $\sigma_{\mathrm{SS}}(\phi) = 0$.

(See Appendix B.1 for a proof).

Also here we can compare the results with those obtained by Lange and Quaas (2007), when the endogenous agglomeration effect is absent. In our analysis, this comes down to posing $d_{\psi}^{(0.5)} = 0$. Three possible outcomes out of the five above presented then arise:[17] *i*) if $d_{\Theta}^{(0.5)} > \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, the symmetric spreading equilibrium is stable independent of trade freeness (case SS–*v*); *ii*) if $\zeta\left(d_{\psi}^{(0.5)}\right) < d_{\Theta}^{(0.5)} < \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, two "break points" emerge $\underline{\phi}_b$ and $\overline{\phi}_b$, and the symmetric spreading is a stable equilibrium only for $\phi \in [0;1] \setminus \left[\underline{\phi}_b;\overline{\phi}_b\right]$ (case SS–*iv*). These two cases correspond to the two possible outcomes mentioned in Proposition 2 in Lange and

[17] Note that cases SS–*i* and SS–*iii* never occur in case of absent market-density effect, since condition $d_{\psi}^{(0.5)} > d_{\psi,0}$ never holds.

Quaas (2007), but with other specific analytical conditions due to the difference in the specification of negative externalities in utility. Adopting a multiplicative formulation as we do generates an additional possible outcome that does not emerge in the analysis by Lange and Quaas (2007), namely: *iii*) if $d_{\Theta}^{(0.5)} < \zeta\left(d_{\psi}^{(0.5)}\right)$, a unique "break point" ϕ_b exists and symmetric spreading is a stable equilibrium only for $\phi \in [0;\phi_b]$ (case SS–*ii*).

1.3 Partial Agglomeration in Symmetric Spatial Configurations

The existence and stability properties of partial agglomeration (PA) of production ($0.5 < h < 1$) can be derived from those of the core-periphery and symmetric-spreading configurations, as summarized in the following proposition.

PROPOSITION 3: Let us consider a ϕ-value such that $0 \leq \phi \leq 1$:

PA–*i*: If the core-periphery ($h = 1$) and symmetric spreading ($h = 0.5$) are both stable equilibria for trade barrier ϕ, then a partial agglomeration equilibrium exists which is unstable.

PA–*ii*: If the core-periphery ($h = 1$) and symmetric spreading ($h = 0.5$) are both unstable equilibria for trade barrier ϕ, then a partial agglomeration equilibrium exists which is stable.

(See Appendix B.2 for a proof).

The emergence and nature (stability vs. instability) of partial agglomeration then depends on the values of $\Theta(1)\psi(1)^{1-\varepsilon}$, $d_{\psi}^{(0.5)}$ and $d_{\Theta}^{(0.5)}$, which determine the stability ranges of the core-periphery (cases CP–*i* to CP–*iii*) and the symmetric-spreading equilibria (*SS–i* to *SS–v*), as well as the relative values of "sustain points" and "break points", whenever these exist.

We draw attention to the case in which CP–*i* and SS–*i* are simultaneously satisfied. This results in the instability of both core-periphery and symmetric-spreading equilibria for all trade barriers. According to Proposition 3, this is associated with the stability of partial agglomeration equilibria for all trade barriers ϕ, which represents a continuous asymmetric distributions of the manufacturing sector across regions. Lange and Quaas (2007) already presented an extension of the basic model by Forslid and Ottaviano (2003) that allows for the existence of some stable partial agglomeration equilibria. However, contrary to previous studies in which the existence of such equilibria is always bound to limited ranges of trade barriers, our framework enables stable partial agglomeration equilibria to emerge for all trade barriers. So our model is capable of explaining a wider range of realistic spatial distributions of population and economic activities. What is more, this property makes our framework valuable for policy analysis and overcomes the shortcomings of previous NEG studies, which have seen very little application to policy.

2- Equilibrium of Non-Symmetric Spatial Configurations.

In this sub-section, we extend the analysis of long term equilibria of the two-region economy to the case of non symmetric configurations characterized by *ex-ante* differences in terms of spatial structure modeled by assuming distinct β-values in the production function: $\beta_1 \neq \beta_2$ (see eq. (4)). We introduce $\nu = \left(\frac{\beta_2}{\beta_1}\right)^{1-\varepsilon}$ and, without loss of generality, assume that condition $\nu < 1$ (corresponding to $\beta_1 < \beta_2$) holds.[18] We then investigate the stability conditions of core-periphery ($h = 0$; $h = 1$) and partial agglomeration ($0 < h < 1$).[19]

2.1 Core-Periphery pattern in non-symmetric configurations

Similar to the analysis carried out in section 1, we first investigate the conditions under which a full agglomeration of production is a stable equilibrium.

According to (18-*b*), the full agglomerations $h = 1$ is stable equilibrium under the condition $\Omega(1,\phi) \geq 0$. Using (19) and (20) shows that this condition is equivalent to $\sigma_{\mathrm{CP}}(\phi) < \frac{1}{\nu}\Theta(1)\psi(1)^{1-\varepsilon}$.

CONDITION 3: The core-periphery pattern $h = 1$ is a stable equilibrium for trade barrier $\phi \in [0;1]$ if: $\frac{1}{\nu}\Theta(1)\psi(1)^{1-\varepsilon} > \sigma_{\mathrm{CP}}(\phi)$.

We can derive the stability conditions of the core-periphery pattern $h = 1$ through a formal analogy with condition 1 by substituting $\Theta(1)\psi(1)^{1-\varepsilon}$ with $\frac{1}{\nu}\Theta(1)\psi(1)^{1-\varepsilon}$. Similarly to

[18] Since we consider regions with different spatial structures, the indirect utility differential $\Omega(h,\phi)$ takes the general form as given in (19). This holds throughout the subsection.

[19] The symmetric-spreading distribution $h = 0.5$ is not discussed in the context of asymmetric configurations ($\beta_1 \neq \beta_2$). The reason is that $\beta_1 \neq \beta_2$ gives $\Omega(0.5,\phi) \neq 0$ (eq. (20)) and this case is not an equilibrium.

Proposition 1, we obtain three possible outcomes for the core-periphery pattern $h=1$ conditional on the position of $\frac{1}{\nu}\Theta(1)\psi(1)^{1-\varepsilon}$ with respect to $\sigma_{\text{CP}}^{\min}$ and 1: *i)* $h=1$ is never an equilibrium if $\frac{1}{\nu}\Theta(1)\psi(1)^{1-\varepsilon}<\sigma_{\text{CP}}^{\min}$; *ii)* $h=1$ is an equilibrium for intermediate trade freeness $\phi\in\left[\underline{\phi}_S^*;\overline{\phi}_S^*\right]$; *iii)* $h=1$ is an equilibrium for low trade barrier $\phi\in\left[\phi_S^*;1\right]$ if $\sigma_{\text{CP}}^{\min}<\frac{1}{\nu}\Theta(1)\psi(1)^{1-\varepsilon}$.[20]

Similarly we derive the stability condition for the full agglomeration $h=0$. According to (18-*c*), the case $h=0$ is a stable equilibrium as long as condition $\Omega(0,\phi)\leq 0$ holds. Using (19) and (20) shows that this condition is equivalent to $\sigma_{\text{CP}}\left(\phi\right)<\nu\Theta(1)\psi(1)^{1-\varepsilon}$.

CONDITION 3BIS: The core-periphery pattern $h=0$ is a stable equilibrium for trade barrier $\phi\in\left[0;1\right]$ if: $\nu\Theta(1)\psi(1)^{1-\varepsilon}>\sigma_{\text{CP}}\left(\phi\right)$.

Similarly to Proposition 1, we obtain three possible outcomes for the core-periphery pattern $h=0$ conditional on the position of $\nu\Theta(1)\psi(1)^{1-\varepsilon}$ with respect to $\sigma_{\text{CP}}^{\min}$ and 1: *i)* $h=0$ is never an equilibrium if $\nu\Theta(1)\psi(1)^{1-\varepsilon}<\sigma_{\text{CP}}^{\min}$; *ii)* $h=0$ is an equilibrium for intermediate trade freeness $\phi\in\left[\underline{\phi}_S^{**};\overline{\phi}_S^{**}\right]$ if $\sigma_{\text{CP}}^{\min}<\nu\Theta(1)\psi(1)^{1-\varepsilon}<1$; *iii)* $h=0$ is an equilibrium for low trade barrier $\phi\in\left[\phi_S^{**};1\right]$ if $\nu\Theta(1)\psi(1)^{1-\varepsilon}>1$.[21]

Two characteristics of the full-agglomeration equilibria are worth noting. First, whatever the ν-value, a full agglomeration is *never* an equilibrium for $\phi=0$ ("no black hole" condition). Second, by assuming $\nu<1$, condition $\Theta(1)\psi(1)^{1-\varepsilon}>\nu\sigma_{\text{CP}}\left(\phi\right)$ is less stringent than condition $\Theta(1)\psi(1)^{1-\varepsilon}>\frac{1}{\nu}\sigma_{\text{CP}}\left(\phi\right)$, so that the stability range of $h=1$ (indicating full agglomeration in region 1) is wider than the stability range of $h=0$ (indicating full

[20] The threshold point ϕ^*_S (with $\underline{\phi}_S^*$ and $\overline{\phi}_S^*$ indicating its upper and lower value in case of existence of multiple points) is the "sustain point" in the case of full agglomeration $h=1$ in non-symmetric configurations (see Proposition 1). It is implicitly given by condition $\frac{1}{\nu}\Theta(1)\psi(1)^{1-\varepsilon}=\sigma_{\text{CP}}(\phi_S^*)$.

[21] The threshold point ϕ_S^{**} (with $\underline{\phi}_S^{**}$ and $\overline{\phi}_S^{**}$ indicating its upper and lower value in case of existence of multiple points) is the "sustain point" in the case of full agglomeration $h=1$ in non-symmetric configurations (see Proposition 1). It is implicitly given by condition $\nu\Theta(1)\psi(1)^{1-\varepsilon}=\sigma_{\text{CP}}(\phi_S^{**})$.

agglomeration in region 2). This makes sense since positing $\nu < 1$ means that region 1 is characterized by a stronger agglomeration effect (as captured by $\beta_1 < \beta_2$), which fosters agglomeration of production.

2.2 Partial Agglomeration in the Non-symmetric Spatial Configuration.

The equilibrium properties of partial agglomeration (PA) of production ($0 < h < 1$) can be derived from those of the core-periphery configurations ($h = 0$ and $h = 1$) as summarized in the following proposition:

PROPOSITION 4: Let us consider a ϕ-value with $0 \le \phi \le 1$:

PA*'–i*: If $h = 0$ and $h = 1$ are both stable equilibria for trade barrier ϕ, then an unstable partial agglomeration equilibrium exists.

PA*'–ii*: If $h = 0$ and $h = 1$ are both unstable equilibria for trade barrier ϕ, then a stable partial agglomeration equilibrium exists.

(See Appendix B.3 for a proof).

According to Proposition 4, if both $h = 1$ and $h = 0$ are never stable equilibria, a stable partial agglomeration equilibrium arises for all trade barrier values. This situation occurs when $\Theta(1)\psi(1)^{1-\varepsilon} \le \nu\sigma_{\mathrm{CP}}^{\min}$.

3- Analysis of the Interplay between Agglomeration and Environmental Effects

The analytical equilibrium conditions derived in the two previous sections complete the earlier NEG literature. To the standard analysis of agglomeration versus dispersion drivers of the equilibrium in the long run[22] we add a new force of agglomeration that stems from an explicit representation of the spatial dimension of the market (which we call 'market density') and take the next step to study it in interaction with an associated environmental externality effect. In our model the 'market density' effect acts as a centripetal force fostering agglomeration of production through the decrease of unitary intra-industry transport costs. We henceforth refer to it as the more general 'agglomeration effect', but we actually mean the specific effect of the size of the market in a given space on the structure of production (see the discussion of equation (4)). On the other hand, the environmental effect acts as a centrifugal force fostering dispersion of activities in response to the negative impact of pollution on

[22] Agglomeration forces, as traditionally approached by the NEG literature are: "market-size" and "cost-of-living" effects, whereas the dispersion force is represented by the "market-crowding" effect [see Forslid and Ottaviano (2003) for a thorough discussion].

domestic utility. The impact of those two effects on long-term equilibria depends on their relative intensities, as defined by functions ψ and Θ, respectively.[23]

Let us start with the stability range of the core-periphery pattern, which can be derived as the ϕ-values for which Condition 1 is satisfied. Everything else being equal, a stronger 'market density' effect, as captured by a lower $\psi(1)$, favors agglomeration by widening the range of trade freeness compatible with stability of the core-periphery,[24] whereas a stronger environmental effect as captured by a lower $\Theta(1)$ fosters dispersion by narrowing this range. The three cases discussed in Proposition 1 differ in terms of the number of emerging "sustain points" resulting from the interplay between the agglomeration and environmental effects. In the case of CP–*i*, there is no "sustain point": the centrifugal environmental effect is so strong that it always renders full agglomeration unstable. In case CP–*ii*, two "sustain points" $\underline{\phi}_S$ and $\overline{\phi}_S$ exist. This corresponds to an intermediate situation in which centrifugal and centripetal forces associated with market density and environmental effects, respectively, are of the same order of magnitude and offset each other. The boundary case $\Theta(1)\psi(1)^{1-\varepsilon}=1$ corresponds to a situation in which the centrifugal environmental and the centripetal market density forces fully offset each other, and lead to an identical result as in the traditional "footloose entrepreneur" model where they are absent: the "sustain point" ϕ_S is unique and is defined by condition $\sigma_{\mathrm{CP}}(\phi)=1$, which is identical to condition (25) in Forslid and Ottaviano (2003). Finally, case CP–*iii* is associated with a unique "sustain point": the centripetal impact of the agglomeration effect dominates, making core-periphery a stable equilibrium in the case of free trade $\phi=1$, so that the range of stability is given by $\phi\in[\phi_S;1]$.

The stability of the symmetric spreading equilibrium is obtained for those ϕ-values for which Condition 2 is satisfied. As formalized in the decomposition of $\sigma_{\mathrm{SS}}(\phi)$ in (25), in determining the stability of the equilibria, our model accounts for the relative importance of agglomeration and environmental effects, as well as for their dependence on the long-run trade costs. Given the relevance and the novelty of this analysis we take the next step of analytically deriving the conditions under which both the interplaying and the trade-dependence mechanisms occur. Due to limited space, we summarize here the main insights from this analysis; mathematical details are given in Appendix B.4.

[23] For the sake of clarity of interpretation of the results, we abstract from considering *ex-ante* differences among regions as in non-symmetric configurations, which, although affecting the results, do not modify the qualitative effects at play.

[24] We recall that $1-\varepsilon<0$, so that a lower $\psi(1)$ means a higher $\psi(1)^{1-\varepsilon}$.

The first term on the right-hand side of (25), ${}^{(\mathrm{FE})}\sigma_{\mathrm{SS}}(\phi)$, captures the forces at play in the traditional footloose-entrepreneur model, as developed by Forslid and Ottaviano (2003), in the absence of any agglomeration and environmental effects ($d_{\psi}^{(0.5)} = d_{\Theta}^{(0.5)} = 0$). In this case, the stability range of the symmetric-spreading equilibrium is determined by condition ${}^{(\mathrm{FE})}\sigma_{\mathrm{SS}}(\phi) < 0$. With (26), this means $\phi < \dfrac{(\varepsilon-\delta)(\varepsilon-1-\delta)}{(\varepsilon+\delta)(\varepsilon-1+\delta)}$, an identical expression to condition (26) obtained by Forslid and Ottaviano (2003).

The last two terms on the right-hand side of (25) describe how agglomeration and environmental effects influence the stability of the symmetric-spreading equilibrium. The second (third) term is positive (negative) [see eq. (26)], so that the agglomeration (environmental) effect unequivocally contributes to the instability (stability) of the symmetric-spreading configuration. A more intense agglomeration (environmental) effect, as captured by a higher $d_{\psi}^{(0.5)}$ ($d_{\Theta}^{(0.5)}$), results in a narrowing (widening) of the stability range of the symmetric-spreading equilibrium.

Proposition 2 provides explicit conditions for the intensity of the agglomeration and environmental effects $d_{\psi}^{(0.5)}$ and $d_{\Theta}^{(0.5)}$, which can be interpreted in terms of the strength of these effects. The high value of $d_{\psi}^{(0.5)}$ in condition $d_{\psi}^{(0.5)} > d_{\psi,0}$ can be viewed as a 'strong agglomeration effect'. Next, $d_{\Theta}^{(0.5)} > \zeta\left(d_{\psi}^{(0.5)}\right)$ means a strong centrifugal force, which we will refer to as a 'strong environmental effect'. Moreover, since $\zeta\left(d_{\psi}^{(0.5)}\right) < \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, condition $d_{\Theta}^{(0.5)} > \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$ is more restrictive than $d_{\Theta}^{(0.5)} > \zeta\left(d_{\psi}^{(0.5)}\right)$ on the value of $d_{\Theta}^{(0.5)}$. Therefore, we interpret the first condition, $d_{\Theta}^{(0.5)} > \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, as representing a 'very strong environmental effect'.

To better understand how the interplay between the agglomeration and environmental effects influences the existence and nature of equilibria, let us give an economic interpretation of the five conditions in Proposition 2.

We first consider the situation where both agglomeration and environmental effects are weak, as captured by $d_{\psi}^{(0.5)} < d_{\psi,0}$ and $d_{\Theta}^{(0.5)} < \zeta\left(d_{\psi}^{(0.5)}\right)$. This corresponds to case SS–*ii* in Proposition 2. This type of stability condition is similar to the one obtained in the traditional footloose-entrepreneur model (Forslid and Ottaviano, 2003). This similarity makes sense, since the addition of small agglomeration and environmental effects only marginally modifies the stability conditions that hold without these effects.

We then turn to the case of a strong environmental and a weak agglomeration effect, as captured by conditions $d_{\psi}^{(0.5)} < d_{\psi,0}$ and $d_{\Theta}^{(0.5)} > \zeta\left(d_{\psi}^{(0.5)}\right)$. Two sub-cases must be distinguished:

- If condition $\zeta\left(d_{\psi}^{(0.5)}\right) < d_{\Theta}^{(0.5)} < \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$ holds, the environmental effect is strong but not 'very strong', and a range of high ϕ-values (close to 1) appears, for which the centrifugal environmental effect is strong enough to favor stability of the symmetric spreading, contrary to the outcome of the traditional footloose-entrepreneur model. As a result, the symmetric-spreading pattern is stable for high and low trade barriers, while it remains unstable for intermediate trade barriers. This corresponds to case SS–*iv* in Proposition 2;
- If condition $d_{\Theta}^{(0.5)} > \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$ holds, i.e. the environmental effect is 'very strong', the associated centrifugal force dominates and uniformly favors the stability of the symmetric-spreading equilibrium. This corresponds to case SS–*v* in Proposition 2 with the symmetric-spreading equilibrium being stable for all values of the trade barrier.

All cases in which condition $d_{\psi}^{(0.5)} > d_{\psi,0}$ holds correspond to a strong agglomeration effect which fosters the instability of the symmetric-spreading equilibrium. The magnitude of the environmental effect, which tends to counterbalance agglomeration, determines the resulting net equilibrium. If a weak environmental effect is assumed ($d_{\Theta}^{(0.5)} < \zeta\left(d_{\psi}^{(0.5)}\right)$), the agglomeration effect dominates the location decisions of economic agents and leads to the instability of the symmetric spreading under all values of the trade barrier (case SS–*i* in Proposition 2). If, on the contrary, a strong environmental effect is considered ($d_{\Theta}^{(0.5)} > \zeta\left(d_{\psi}^{(0.5)}\right)$), the trade-off between agglomeration and environmental effect dominates the stability conditions of the symmetric-spreading equilibrium. It is found that the positive agglomeration (negative environmental) effect dominates at low (high) ϕ-values, leading to the instability (stability) of the symmetric-spreading equilibrium (see Appendix A.4 for a proof). This corresponds to case SS–*iii* in Proposition 2.

4- Spatial Equilibria

Combining the conditions that define the nature of the long-run equilibria presented in previous sub-sections gives rise to several alternative types of bifurcation diagrams. For the sake of realistically representing the spatial organization of the world economy, we exclude in the remainder of the paper the extreme equilibria that one would not expect to find in practice. In particular, we limit the analysis to equilibria other than full agglomeration, and instead focus on those equilibria that allow for a (stable) partial agglomeration to arise. To achieve this, we adopt adequate numerical values of model parameters and exogenous variables, as well as the functional specifications, whose rationale is discussed in Appendices A.5 and A.6, respectively.

We start by considering stable (continuous curves) and unstable (broken curves) equilibria in the symmetric configurations A and B, as defined by condition (18) in the case $\beta_1 = \beta_2$ (see Figure 4.1). This is the typical 'spatial equilibrium' picture that arises in "footloose entrepreneur"-like models, given the absence of an explicit spatial dimension here.

Figure 4.1: *Long-run Equilibria for Symmetric Spatial Configurations.*

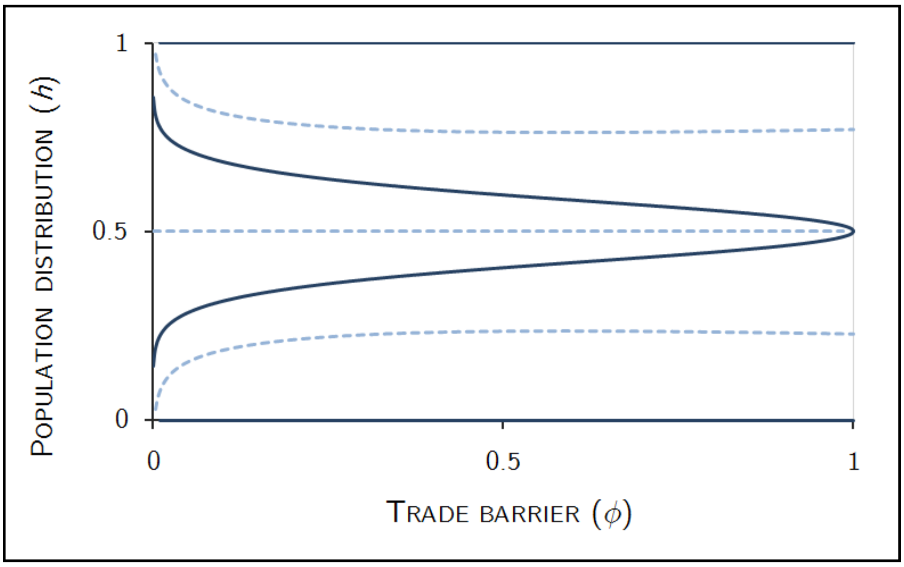
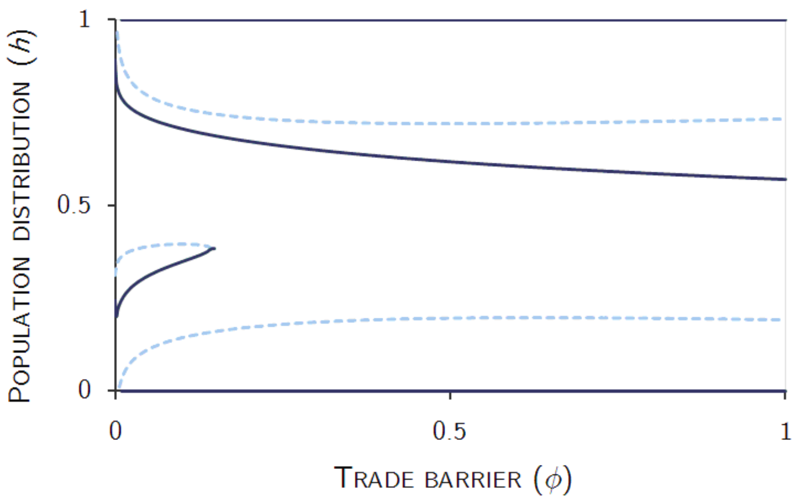


Figure 4.2 displays long-run stable (continuous curve) and unstable (broken curve) equilibria for the asymmetric configuration C, as defined by condition (18) in the case $\beta_1 \neq \beta_2$. Again, partial agglomeration equilibria emerge for any trade barrier. However, unlike in the previous case, the partial agglomeration equilibria are not symmetric around $h = 0.5$, as a consequence of *ex-ante* differences between the two regions in the spatial

setting. More precisely, a partial agglomeration in the (urbanized) region 1 ($0.5 < h < 1$) can be a long-run equilibrium for any trade barrier, while a partial agglomeration in the (undeveloped) region 2 ($0 < h < 0.5$) can be a long-run equilibrium only for sufficiently high trade barriers. Firms may indeed agglomerate in the undeveloped region only if high trade costs ensure a strong incentive for a relocation of production close to consumption places.

Figure 4.2: *Long-run Equilibria for the Asymmetric Spatial Configuration*.



Our analysis extends the traditional NEG approach, in that it allows for the emergence of a stable partial agglomeration for any trade barrier. To the best of our knowledge, this constitutes a major novelty in the NEG literature, where partial agglomeration is traditionally either unstable – as in the seminal paper by Krugman (1991) and in Forslid and Ottaviano (2003) when symmetric unskilled labor endowment is assumed – or stable, but only for a limited range of trade barriers (see Forslid and Ottaviano (2003); and Lange and Quaas (2007)).

III. Extension to Spatial Sustainability

In this section we provide the second main extension to the standard NEG modeling framework and address the long-term interplay between trade, pollution and spatial location of economic activities, and its counterpart transport in terms of environmental sustainability. In so doing we touch on the literature that goes under the name of 'trade and environment', and which focuses on the role of trade costs on pollution [Copeland and Taylor (1994; 1995)]. We hence discuss our results in the light of the findings of this mainstream literature.

1. Global Pollution and Trade

In the strand of literature that followed Copeland and Taylor (1994; 1995), the relation between spatial concentration and environmental externalities is mainly driven by the negative effect of pollution on utility and productivity (Copeland and Taylor, 1999); Benarroch and Thille, 2001; Unteroberdoerster, 2001). We engage with these analyses by considering *i*) the role of the spatial dimension of the economy on pollution through the impact of the type of spatial structure (urbanized versus undeveloped land use) on intra-regional transport requirements, as captured by parameter β in eq. (4); *ii*) the positive impact of agglomeration on productivity through a decrease of unitary energy requirements for production in more agglomerated production patterns as captured by variable $\bar{\psi}\left(n_j\right)$ in (4).

For the purpose of analyzing environmental sustainability in the context of the climate change debate (and other environmental issues like depletion of the ozone layer or biodiversity), we consider global pollution affecting equally both regions, regardless of its source. Hence, we generalize the definition of environmental externalities in (13) to include transboundary pollution as generated by international trade.[25] We assume trade-associated pollution to be proportional to the volume of trade that is necessary to satisfy the domestic demand for imported goods. Quantities Tc_{jk} and Tc_{kj} represent such volume for each of the good variety that is produced in region *j* and *k* respectively, with $j \neq k$. Introducing a^G and b^G as the intensity of global pollution generated by manufacturing production and interregional trade/transport respectively, global pollution flow E^G is given by:[26]

$$E^G = a^G \left(\xi_j n_j x_j + \xi_k n_k x_k \right) + b^G \left(Tc_{jk} n_j + Tc_{kj} n_k \right). \tag{27}$$

The amount of long-term global emissions in (27) depends on trade freeness through the interplay between three effects. First, the iceberg structure for transport costs implies that lower trade barrier brings about a decrease of emissions per unit of good shipped. Second, for a given spatial distribution of the economic activities, a trade barrier affects the quantity of good actually shipped from one region to another, since freer trade fosters a higher demand for imported good in both regions. This ultimately acts in the direction of increasing trade-

[25] Extending the analysis in Sections I and II to consider trade-related global pollution emissions does alter neither the nature nor the quality of the equilibrium findings, as these emissions play no role in individuals' migration decisions (and are hence ignored in the analysis of long-run equilibria).

[26] Note that regardless of the source of emission, pollution has both a local and global component of impact. In the case of manufacturing production, we can imagine pollution resembling particulate matter (local impact) and carbon emission (global impact).

related emissions. Third, trade freeness affects the long-term location of economic activities (see Figures 4.1 and 4.2), which ultimately determines both the intensity of trade as a consequence of the uneven distribution of production and consumption locations[27], and the level of global economic activity because of the effects of agglomeration on regional and aggregate production.

2. Pollution Dynamics and Assimilation Potential

For the sake of comparability with the literature on trade and the environment following Copeland and Taylor (1999), we consider the level of environmental capital stock K as deprived or enhanced over time, depending on the flow of pollution E^G and the "nature's regenerative capacity" $F(K)$, as described in the standard dynamic equation:

$$\frac{dK}{dt} = F(K) - E^G. \tag{28}$$

In the literature following Copeland and Taylor (1999), the function $F(K)$ is traditionally taken as a linear function of the difference with a "natural level" $\bar{K}$ at a "recovery rate" g: $F(K) = g\left(\bar{K} - K\right)$ (see for example, eq. (1) in Copeland and Taylor (1999)). With this specification, environmental capital gravitates towards the "natural level" $\bar{K}$ in case of absent environmental pollution, and $\frac{dK}{dt}$ is still positive at $K = 0$, which means that no minimum viable level of environment's capital stock occurs. This result is possible due to the underlying assumption that *substitutability* occurs between environmental and man-made capital, which is evidently an oversimplification of the description of pollution dynamics.

Unlike Copeland and Taylor (1999) we consider that such a minimum exists as a consequence of negative nonlinear mechanisms limiting the "nature's regenerative capacity" at low levels of environmental capital.[28] We represent the existence of these mechanisms by a drop in the "nature's regenerative capacity" $F(K)$ at low levels of environmental capital stock. To keep the analysis simple we distinguish between two regimes of "nature's regenerative capacity" and assume that: *i*) at medium and high levels of environmental stock,

[27] For example, full agglomeration in one region creates the need to export manufacturing goods for unskilled workers in the other region.

[28] This is particularly true when one considers the case of climate change. Here non-linear events include ocean acidification at high CO_2 atmospheric concentrations, which may cause a drastic drop of biological carbon sequestration in oceans and affect significantly the "regenerative capacity" of the climate.

the “nature’s regenerative capacity” is a constant process A (Gradus and Smulders, 1993); Keeler et al, 1971; van der Ploeg and Withagen, 1991], which we call the ‘pollution assimilation potential’.[29] This represents thus the maximum of the natural recovery rate of the environment, like the (natural) biological and ocean carbon sequestration in the case of climate change, or the rate of reproduction of (rare) biological species; *ii*) at a low level of the environmental stock, this “nature’s regenerative capacity” drops to zero. This leads to:

$$F(K) = \begin{Bmatrix} A & if\ K \geq \bar{K} \\ 0 & if\ K < \bar{K} \end{Bmatrix}. \tag{29}$$

Here, $\bar{K}$ is a threshold below which the regenerative capacity drops to zero. The functional specification for the “nature’s regenerative capacity” ensures that $\frac{dK}{dt} < 0$ at $K = 0$, or that falling below a certain level of environmental quality may turn out irreversible. This leads us to adopt a more restrictive approach to sustainability than in the traditional approach assuming *substitutability* between natural and man-made capital. Instead, we acknowledge the specificity or unique character of the environment by imposing to maintain its stock because the functions it performs cannot be duplicated by manufactured capital. This assumption seems particularly relevant for the large-scale environmental issues which motivate this analysis (climate, ozone layer, biodiversity) and means positing $\frac{dK}{dt} \geq 0$ in the long-run equilibrium, which according to (28) and (29) leads to:

$$E^G(h, \phi) < A. \tag{30}$$

3. Formalizing Spatial Sustainability

Here, we address the spatial or geographical nature of sustainability, in the sense defined in the previous subsection. This means investigating the conditions under which the long-run equilibrium conditions of the two-region economy analyzed in Section II are compatible with condition (30). A sustainable long-run equilibrium exists if one of the three following sets of conditions is satisfied:

[29] Our (simplified) specification of the dynamics of pollution for high levels of the environmental capital stock is motivated by environmental issues like climate change, which involves irreversibility and catastrophic mechanisms that compromise the “nature’s regenerative capacity” at low levels of the environmental stock.

$$a)\begin{cases}0<h<1\\ \Omega(h,\phi)=0, \dfrac{\partial\Omega}{\partial h}(h,\phi)<0;\\ E^s(h,\phi)\le A\end{cases}\quad b)\begin{cases}h=1\\ \Omega(h,\phi)\ge 0\\ E^s(h,\phi)\le A\end{cases};\quad c)\begin{cases}h=0\\ \Omega(h,\phi)\le 0\\ E^s(h,\phi)\le A\end{cases}. \tag{31}$$

If condition *a*) in (31) holds, the distribution of population with a share *h* of workers in region 1 (and a share $(1-h)$ in region 2) is a sustainable long-run equilibrium for the spatial configuration considered. If condition *b*) (condition *c*)) holds, a full agglomeration of skilled workers in region 1 (region 2) is the sustainable long-run equilibrium. Finally, if none of the three conditions is satisfied, the spatial configuration is always unsustainable, that is, for all possible trade barriers and spatial distributions of the population.

For a given configuration λ, we define $E_\lambda^{G,\min}$ and $E_\lambda^{G,\max}$ to denote the minimum and maximum long-run levels of polluting emissions over all ranges of trade barrier. Three cases can occur:

CASE 1: For high configurations in terms of emission characterized by the condition $E_\lambda^{G,\min}>A$, the level of long-run pollution is always larger than the pollution assimilation potential *A*, whatever the trade barrier ϕ. This means that the spatial configuration considered is always unsustainable.

For any other type of configuration, there exists a range of ϕ-values that satisfies (30). We define $\phi_\lambda^*(A)$ as the highest of these values (lowest trade barrier), meaning the minimum constraint on the economy in terms of barriers to trade to ensure a sustainable development of a spatial configuration λ.[30] Setting a constraint on long-run emissions, as imposed by spatial sustainability in (30) results from the interplay between minimum trade barriers, $\phi_\lambda^*(A)$, the λ-specific pollution stock, and the pollution assimilation potential under normal conditions, *A*, as follows:

$$E_\lambda^s\left(h,\phi_\lambda^*(A)\right)=\min\left(A,E_\lambda^{s,\max}\right). \tag{32}$$

CASE 2: For intermediate configurations in terms of emission characterized by the condition $E_\lambda^{G,\min}<A<E_\lambda^{G,\max}$, the relation in (32) becomes $E_\lambda^G\left(h,\phi_\lambda^*(A)\right)=A$. The corresponding

[30] Intuitively, condition $\phi_\lambda^*(A)=0$ expresses the unsustainability of the above Case 1.

solution in $\phi_\lambda^*(A)$ is such that: $0<\phi_\lambda^*(A)<1$. In this case, the configuration considered satisfies (does not satisfy) the sustainability conditions on emissions if trade barriers are high (low) enough: $\phi<\phi_\lambda^*(A)$ ($\phi>\phi_\lambda^*(A)$).

CASE 3: For low configurations in terms of emission characterized by the condition $E_\lambda^{G,\max}<A$, the configuration considered satisfies the sustainability conditions on emissions whatever the value of trade barrier. In particular, the configuration remains sustainable even if free trade is assumed as expressed by $\phi_\lambda^*(A)=1$.

We are now able to assess the sustainability characteristics of any spatial configuration λ by considering the minimum value of the trade barrier $\phi_\lambda^*(A)$ that ensures that the stock of environmental capital increases in the long term.

4- Numerical Application

By substituting (2) and (11) into (27), and considering baseline values for the model parameters and exogenous variables (see Appendix A.5), we compute the threshold values for stock pollution $E_\lambda^{s,\min}$ and $E_\lambda^{s,\max}$ for each of the three configurations $\lambda\in\{\mathrm{A,B,C}\}$ (Table 4.1).

Table 4.1 *Threshold values for emissions in the three spatial configurations*

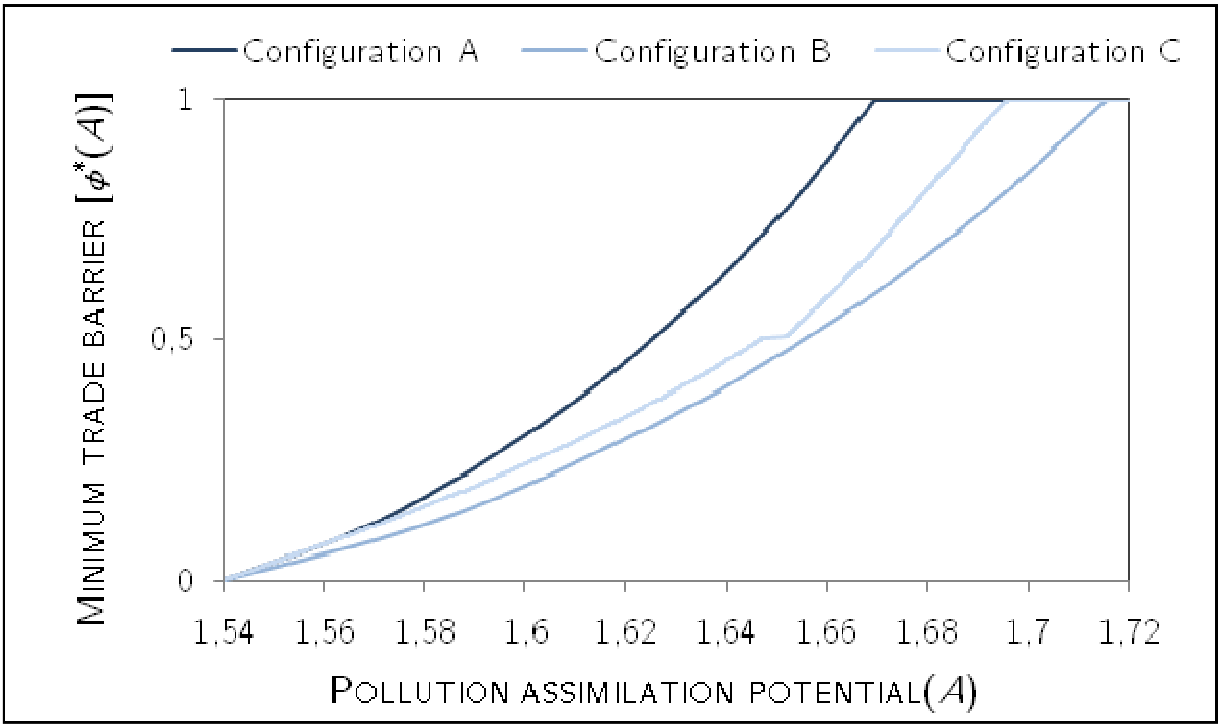
Spatial Configuration	**Emission threshold values**	
	$E_\lambda^{s,\min}$	$E_\lambda^{s,\max}$
A (both regions with undeveloped land)	1.54	1.67
B (both regions with urbanized land)	1.54	1.72
C (one region urbanized, other undeveloped)	1.54	1.70

We then compare the performance of the three spatial configurations in terms of sustainability of the final long-run spatial equilibrium by investigating the variation of the minimum trade barrier for sustainability $\phi_\lambda^*(A)$ across the configurations for a given level of the pollution assimilation potential A.

We limit the analysis to the range of A-values that are not associated with trivial outcomes, that is, we exclude those for which the three spatial configurations under consideration are either never or always sustainable. For this purpose, we define $A^{\min}$ and $A^{\max}$ as the minimum and maximum values of $E_{\lambda}^{G,\min}$ and $E_{\lambda}^{G,\max}$ across all spatial configurations, respectively. In other words, $A^{\min} = \min_{\lambda \in \{A,B,C\}} E_{\lambda}^{s,\min}$ and $A^{\max} = \max_{\lambda \in \{A,B,C\}} E_{\lambda}^{s,\max}$.

For the numerical data summarized in Table I, $A^{\min} = 1.54$ and $A^{\max} = 1.72$. In order to perform a non-trivial analysis of the sustainability of the configurations, we do not consider the cases where $A < A^{\min}$ with the three configurations in Case 1, nor $A > A^{\max}$ with all configurations in Case 3. In other words, we only consider values of A satisfying $A^{\min} \leq A \leq A^{\max}$ and are interested in the numerical values of the lowest trade barrier that satisfies the sustainability condition in (30) for all the three spatial configurations $\lambda \in \{\text{A}, \text{B}, \text{C}\}$. Figure 4.3 summarizes graphically the results.

FIGURE 4.3: *Minimum Trade Barriers Assuring Sustainability of the Three Spatial Configurations*



For each configuration $\lambda \in \{\text{A,B,C}\}$, the stringency of the barriers to trade that are necessary to achieve sustainability is captured by the value of $\phi_{\lambda}^{*}(A)$: the higher $\phi_{\lambda}^{*}(A)$, the wider the range of trade barrier values that is compatible with sustainability for a given pollution assimilation potential *A* and configuration λ.[31]

Figure 4.3 shows that conditional to the spatial configuration considered, free trade (defined by $\phi = 1$) can be incompatible with environmental sustainability. The reason for this is that global emissions of pollutants reach their maximum $E_{\lambda}^{s,\max}$ for completely lax trade restrictions. When the emission level $E_{\lambda}^{s,\max}$ is higher than the pollution assimilation potential *A*, a more stringent trade barrier is required to meet the sustainability condition in (30), as captured by $\phi_{\lambda}^{*}(A) < 1$.

For a given value of the pollution assimilation potential *A*, the differences in values of $\phi_{\lambda}^{*}(A)$ across configurations illustrate that the stringency of the trade restrictions depends critically on the spatial configuration considered. In particular, it appears that $\phi_{\lambda}^{*}(A)$ is always higher for configuration A than for the two others, implying that any trade barrier that ensures sustainability of either configuration B or C also assures configuration A is sustainable.

A more detailed analysis requires a study of different *A*-range values over the interval $\left[A^{\min}; A^{\max}\right]$, as summarized in Table 4.2.

Table 4.2: *Sustainability of spatial configurations for different pollution assimilation potentials*

Spatial configuration	**Value range of the pollution assimilation potential *A***		
	$\left[A^{\min} = 1.54; 1.67\right]$	$\left[1.67; 1.70\right]$	$\left[1.70; 1.72 = A^{\max}\right]$
A	$\exists\phi$	$\forall\phi$	$\forall\phi$
B	$\exists\phi$	$\exists\phi$	$\exists\phi$
C	$\exists\phi$	$\exists\phi$	$\forall\phi$

Legend: The symbol '$\exists\phi$'denotes that the condition for sustainability is satisfied only for certain values of the trade barrier parameter (namely, $\phi < \phi_{\lambda}^{*}(A)$, with $\phi_{\lambda}^{*}(A) < 1$ associated with the case $E_{\lambda}^{s,\min} < A < E_{\lambda}^{s,\max}$). The symbol '$\forall\phi$' denotes that the condition for sustainability is *always* satisfied, for any value of the trade barrier parameter (since $\phi_{\lambda}^{*}(A) = 1$, associated with the case $A > E_{\lambda}^{s,\max}$).

In the non-trivial case $A \geq A^{\min}$, several spatial configurations potentially meet the spatial sustainability requirements summarized by condition $\phi < \phi_{\lambda}^{*}(A)$. The associated constraint on trade barriers depends on the spatial configuration considered. For example, for $1.67 \leq A \leq 1.70$, three sustainable equilibria are possible: either production chooses a

[31] We recall that configuration λ is sustainable as long as the trade barrier ϕ satisfies $\phi < \phi_{\lambda}^{*}(A)$.

configuration A-like pattern (both regions feature undeveloped land use) and no constraint on trade barriers is necessary for sustainability purposes (condition $\forall\phi$ in Table 4.2); or the system moves to a B- or C-like spatial configuration, in which case trade barriers must be set high enough to satisfy condition $\phi < \phi_{\lambda}^{*}(A) < 1$. The policy relevance of this result is that it shows the sustainability offsetting effect between a policy imposing barrier to the volume of inter-regional or international trade and a policy inducing a reorganization of the spatial structure of regional economies.

IV. Conclusion

This paper has developed a theoretical approach to study the impact of spatial configurations of the world economy in the presence of local and global environmental externalities and a sustainability condition. The starting point was the notion of spatial sustainability, denoting a spatial configuration that is consistent with long-term pollution being within the pollution assimilation potential. Our framework accounts for three drivers of economic activity and (un)sustainability, namely, agglomeration spillovers, advantages of trade, and environmental externalities. It extends the "footloose entrepreneur" model of the new economic geography by introducing four major innovations. First, it formalizes heterogeneous patterns of land development of the economy in multiple regions (which we call 'spatial configurations'). Second, agglomeration spillover effects emerge endogenously from the analytical setting. In particular, we allow for simultaneous modeling of increasing returns to scale at the firm level and external economies at the industry level. Third, the model generates continuous and asymmetric distributions of population and economic activity in space, hence allowing for non-extreme, realistic spatio-economic configurations associated with a full range of trade barriers (i.e. inherent costs, possibly partly due to trade policies). Fourth, it includes the dynamics of pollution, and distinguishes between environmental externalities and environmental sustainability, which represents an innovation over the existing literatures on geography, trade and environment as well as on the economic analysis of sustainability. A number of additional, minor novelties were introduced that render the application of our framework of potential interest to analysts of environmental policy and sustainable trade. This ultimately contributes to overcoming the limitations of the NEG approach which has traditionally seen little application to policy questions and environmental issues.

The starting point of our analysis was a spatial-economic structure which includes manufacturing and traditional production sectors. Regions are characterized by alternative degrees of land development resulting in potentially different levels of agglomeration spillovers. By formalizing an endogenous agglomeration effect, we capture the intensity of the spatial spillover associated with the regional 'market density', reflected by the number of firms that are active in a given space, and the regional 'market form', reflected by the capacity of infrastructure endowment. Agglomeration affects environmental pollution through two mechanisms. First, agglomeration increases the scale of production activity leading to more energy use and associated pollutive emissions. Second, agglomeration reduces transport requirements for production through a lower intraregional transport intensity of production (shorter distances) and learning and R&D spillovers leading to improved energy technologies (i.e. lower emissions). Because of these opposite mechanisms, the interplay between trade and agglomeration in determining a stable spatial distribution of economic activity and sustainable levels of pollution is not trivial, that is, the net general equilibrium outcome is not obvious ex-ante.

The model extends previous NEG studies by presenting analytical solutions that have an intuitive interpretation of the relative intensity of (a comprehensive set of) equilibrium offsetting agglomeration and dispersion forces over the whole range of trade barriers. This in turn allows for continuous and asymmetric distributions of population and economic activities across space. The sustainability characteristics of the long-run equilibrium have been studied here over the whole range of the 'pollution assimilation potential', which is taken as a measure of the "nature's regenerative capacity", by allowing for variations in the trade barrier and the regional spatial form of the three alternative configurations. A main result of the analysis is that regional concentration of economic activities is not necessarily the most sustainable spatial structure. This is only true for a very lax environmental sustainability constraint which is due to a very high pollution assimilation potential. A counter-intuitive finding is that a dispersion of economic activities may perform best in terms of sustainability for very low values of the pollution assimilation potential because of less stringent trade barriers resulting from the sustainability condition.

The results illustrate the relevance of sustainability-offsetting spatial structure and trade effects in environmental regulation. This insight can help to formulate an effective and efficient mix of policies focusing directly on emissions reduction, redirection of trade or spatial reorganization. More specifically, the approach provides a theoretical basis for further investigation of the interaction and complementarity of such diverse instruments as pollution taxes, technological standards, land taxation, road pricing and parking tariffs. Moreover, since the approach makes a clear distinction between environmental externalities and environmental sustainability, it adds a dynamic element to the existing literature on trade and environment which makes it very suitable for addressing the spatial economic and welfare dimension of long-term environmental problems such as climate change.

REFERENCES

Baldwin R, Martin P, Ottaviano G and Robert-Nicoud F (2003). *Economic Geography and Public Policy*. Princeton University Press.

Benarroch M and Thille H (2001). Transboundary Pollution and the Gains from Trade. J. Int. Econ. 55: 139–159.

Bernard AB, Eaton J, Jensen BJ and Kortum S (2003). Plants and Productivity in International Trade. Amer. Econ. Rev. 93: 1268–1290.

Brakman S, Garretsen H, Gigengack R, van Marrewijk CJ and Wagenvoort R (1996). Negative Feedbacks in the Economy and Industrial Location. J. Reg. Sci. 36: 631–651.

Ciccone A (2002). Agglomeration Effects in Europe. Eur. Econ. *Rev.* 46: 213–227.

Ciccone A and Hall RE (1996). Productivity and the Density of Economic Activity. Amer. Econ. Rev. 86: 54–70.

Copeland BR and Taylor MS (1994). North-South Trade and the Environment. Quart. J. Econ. 109: 755-787.

Copeland BR and Taylor MS (1995). Trade and Transboundary Pollution. Amer. Econ. Rev. 85: 716–737.

Copeland BR and Taylor MS (1999). Trade, Spatial Separation and the Environment. J. Int. Econ. 47: 137–168.

Copeland BR and Taylor MS (2004). Trade, Growth, and the Environment. J. Econ. Lit. 42: 7–71.

Dixit AK and Stiglitz JE (1977). Monopolistic Competition and Optimum Product Diversity. Amer. Econ. Rev. 67: 297–308.

Duranton G and Puga D (2004). Micro-Foundations of Urban Agglomeration Economies. In Henderson,V. and Thisse,J-F. (eds.), *Handbook of Regional and Urban Economics: Cities and Geography*. Elsevier, Amsterdam, NL, pp. 2063–2117.

Ebert U and Welsch H (2004). Meaningful Environmental Indices: A Social Choice Approach. J. Environ. Econ. Manage. 47: 270–283.

Elbers C and Withagen CA (2004). Environmental Policy, Population Dynamics and Agglomeration. The B.E. Journal of Economic Analysis & Policy, 3.

Eppink F and Withagen CA (2009). Spatial Patterns of Biodiversity Conservation in a Multiregional General Equilibrium Model. Resource Energy Econ. 31: 75–88.

Forslid R and Ottaviano GIP (2003). An Analytically Solvable Core-periphery Model. J. Econ. Geog. 3: 229–240.

Fujita M and Thisse JF (2002). *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth.* Cambridge University Press, Cambridge, UK.

Fujita M, Krugman PR and Venables AJ (1999). *The Spatial Economy: Cities, Regions, and International Trade.* MIT Press, Cambridge MA.

Gradus R and Smulders S (1993). The Trade-off between Environmental Care and Long-Term Growth-Pollution in Three Prototype Growth Models. J. Econ. 58: 25–51.

Grazi F, van den Bergh JCJM and Rietveld P (2007). Spatial Welfare Economics versus Ecological Footprint: Modeling Agglomeration, Externalities, and Trade. Environ. Resource Econ. 38: 135–153.

IEA. (2008). *World Energy Outlook 2008*. International Energy Agency, IEA/OECD, Paris.

IPCC. (2007). *Climate Change 2007: The Physical Science Basis. Summary for Policymakers*. Intergovernmental Panel on Climate Change, IPCC, Geneva, SW.

Jeppesen T, List JA and Folmer H (2002). Environmental Regulations and New Plant Location Decisions: Evidence from a Meta-analysis. J. Reg. Sci. 42: 19–49.

Keeler E, Spence M and Zeckhauser R (1971). The Optimal Control of Pollution. J. Econ. Theory 4: 19–34.

Keller W (2002). Geographic Localization of International Technology Diffusion. Amer. Econ. Rev. 92: 120–142.

Krugman PR (1991). Increasing Returns and Economic Geography. J. Polit. Economy 99: 483–499.

Lange A and Quaas MF (2007). Economic Geography and the Effect of Environmental Pollution on Agglomeration. The B.E. Journal of Economic Analysis & Policy, 7.

Martin P and Rogers CA (1995). Industrial Location and Public Infrastructure. J. Int. Econ. 39: 335–351.

OECD. (2007). *2007-2008 Programme of Work of the OECD Territorial Development Policy Committee*. Organization for Economic Co-operation and Development, OECD, Paris.

OECD. (2010). *Regional Development Policies in OECD Countries.* Organization for Economic Co-operation and Development, OECD, Paris.

Ottaviano GIP (2003). Regional Policy in the Global Economy: Insights from New Economic Geography. Reg. Stud. 37: 665–673.

Ottaviano GIP and Thisse JF (2004). Agglomeration and Economic Geography. In Henderson,V. and Thisse,J-F. (eds.), *Handbook of Regional and Urban Economics: Cities and Geography*. Elsevier, Amsterdam, NL, pp. 2563–2608.

Pezzey JCV and Toman MA (2005). Sustainability and its Economic Interpretations.” In Simpson,R.D., Toman,M.A. and Ayres,R.U. (eds.) *Scarcity and Growth: Natural Resources and the Environment in the New Millennium* RFF Press, Washington, D.C, pp 121–142

Rauscher M (2003). Hot Spots, High Smoke Stacks, and the Geography of Pollution. Paper presented at the workshop Spatial Environmental Economics, University of Heidelberg.

Samuelson PA (1952). The Transfer Problem and Transport Costs: The Terms of Trade when the Impediments are Absent. Econ. J. 62: 278–304.

Scitovsky T (1954). Two Concepts of External Economies. J. Polit. Economy 62: 143–151.

Unteroberdoerster O (2001). Trade and Transboundary Pollution: Spatial Separation Reconsidered. J. Environ. Econ. Manage. 41: 269–285.

van der Ploeg F and Withagen CA (1991). Pollution Control and the Ramsey Problem. Environ. Resource Econ. 1: 215–236.

van Marrewijk,C. (2005). Geographical Economics and the Role of Pollution on Location. Tinbergen Institute Tinbergen Institute Discussion Papers.

Zeng,D-Z. and Zhao,L. (2006). Pollution Havens and Industrial Agglomeration. Research Institute for Economics & Business Administration, Kobe University Discussion Paper Series 197.

Chapter 5

City dynamics under macroeconomic constraints

Chapter 5[*] extends the theoretical approach of agglomeration and environmental externalities developed in the previous chapter to improve its robustness with respect to empirical evidence. The model provides an improved representation of the spatial dimension of the economy in which multiple cities potentially emerge and interactions occur among them and between cities and the surrounding rural area. The breakdown of the OECD economy in a mass of urban areas is enabled through calibration of the 74 largest agglomerations according to the OECD metropolitan dataset. The dynamics is driven by changes in the macroeconomic conditions and firms' migration decisions towards the most attractive locations; it is tested against historical data, and proves to reproduce well past trends. We further analyze its characteristics as a forecasting tool by investigating the mechanisms and the major determinants of the dynamic process.

[*] This chapter is the result of a joint work with Fabio Grazi.

The patterns along which an economy locates and the determinants driving its distribution across the space are the core subject of the new economic geography (NEG) developed after Krugman (1991). Comprehension of the spatial determinants of regional economic development is set at the core issue of NEG's investigation, which particularly looks at how firms and households agglomerate or sprawl. The NEG approach makes use of a set of assumptions that combine (a) general equilibrium modeling, (b) increasing returns related to indivisibilities at the firms' level and monopolistic competition à la Dixit and Stiglitz (1977), (c) transport/trade costs (traditionally modeled in the 'iceberg' form *à la* Samuelson (1952), which make location matter, and (d) the movement of production factors and agents.

Yet, NEG frameworks traditionally fall short of rendering a picture of the spatial organization and distribution of activities that can be observed in the real world. This lack of realism is at the origins of the lack of implications by NEG that are meaningful to policy makers. In order for NEG approaches to achieve a certain degree of empirical validity that may translate in sharp policy conclusions, some important extensions need to be addressed, as Fujita and Mori (2005) already pointed out in their 'way forward' for second-generation NEG models. First, unifying Urban economics and the New Economic Geography is necessary to investigate consistently the development of cities (having spatial extent) and industrial locations in the same continuous space, the former providing insights on the intra-city structure while the latter informs on the spatial distribution among cities. Second, the objective of explicit empirical validity imposes to go beyond analytical approaches valid in two-region-two-sector models to adopt numerically computable results in the context of multiple available (urban) locations. Bahrens and Nicoud have pushed this critique forward in their article prepared for the 10th anniversary issue of *the Journal of Economic Geography* and called for NEG necessarily opening up to calibration excercises and quantified models. To date, none of these extensions are available.

Against this backdrop, we address the two above identified extensions of standard NEG models by designing a computable general equilibrium framework that allows for a) multiple cities in mutual interaction, whose number and size is endogenously determined; and b) dynamic recursive representation of them under the constraint of global macroeconomic trends. To the best of our knowledge, the spatial breakdown of the firms' location preferences over multiple endogenous agglomerations and simulations of their dynamic properties over the long term are simply not available in the NEG framework.

In order to render a realistic picture of spatial distribution patterns of an economy across a set of available urban agglomerations, we represent intra-urban benefits and costs that are both function of how close to the city-center the households' location preference falls, in line with standard approaches in Urban Economics. This provides a description of intra-regional spatial structure. We assume that all firms have identical fixed production costs structures but face different costs according to the specificity of their location among a number of agglomerations. Our approach distinguishes agglomeration-specific benefits in the form of economies of scale that arise from economic agents being located close together (e.g., facilitation of interchange among firms, job flexibility, ability to support social and cultural events). On the other hand, we also consider agglomeration-specific costs due to diseconomies that arise from congestion, like increased land costs and higher commuting costs for workers.

Dynamics of the urban system is enabled through three levels of interaction at the local, regional and national scales between intra-urban land-use patterns, migration decisions of firms (whose location preferences go towards those agglomeration markets that offer the best investment opportunities) and the underlying constraint of macroeconomic trends at the national level, respectively.

Section I presents the static model and section II extends it to dynamics. The model is calibrated on empirical data for 74 OECD agglomerations in Section III, and its validity to reproduce past trends is verified in Section IV. Section V investigates the long-run equilibrium and extends the standard 'bifurcation diagram' to a multiplicity of 'regions'. Finally, Section VI analyses the determinants of the dynamics through a long-term forecasting exercise for urban population at 2050 horizon.

I- The static space economy

A (group of) country(ies) is envisaged as a mass of N_A+1 regions, with N_A urban agglomerations and a "rural area".[1] In the former, land is conceived as a heterogenous space for households and firms produce a number of varieties i of a differentiated (manufacturing) good M under increasing return to scale. In the rural area, land is conceived as a homogenous

[1] The term "rural" is used for the sake of conveniency by opposition to the agglomerations. But, this does not mean that is an agricultural-type of production. Instead, it recovers all economic activity that is not realized in the N_A agglomerations, including industrial production and some manufacturing and service activities.

space, the L_A households are strictly identical and production is made of a homogenous good *f* under constant returns to scale.

1. The urban economies

In each agglomeration $j \in [1; N_A]$, three types of agents are operating: n_j firms[2], L_j households and a local government.

1.1 Firms

Manufacturing production uses capital and labor as its two spatially mobile input factors. Production costs differ across agglomerations as a result of heterogeneous labor productivity ($l_j \neq l_k, \forall \{j,k\} \in [1 \;\; N_A]$), they are identical for all firms of a given agglomeration *j*.[3]

Labor is the variable factor of production and is subject to external economies of scale so that unitary labor costs LC_j are reduced in a larger market, as follows:

$$LC_j = \frac{l_j w_j}{n_j^{\alpha}} \tag{1}$$

$\alpha > 0$ is the elasticity of labor costs to market size, measured by the number of active firms in region *j*. It captures the improvement of effective productivity permitted by the agglomeration of production through facilitated technology spillover.

Capital is the fixed factor of production, and, with fixed input requirement χ, the amount X_j of productive capital in agglomeration *j* is proportional to the number of domestic firms, n_j:

$$X_j = \chi n_j \tag{2}$$

Letting r_j be the unitary return of capital X_j, the total cost TC_j of producing q_j for a firm settled in agglomeration *j* is expressed as:

---

[2] Increasing returns foster the concentration of production of each variety in a single firm so that the number of firms n_j that are active in agglomeration *j* represents the number of varieties produced there.

[3] This means that all varieties produced here are identical in terms of prices and quantities and allows us to drop the notation *i* for the variety in the remainder of the analysis.

$$TC_j = \chi r_j + \frac{l_j w_j}{n_j^{\alpha}} q_j \tag{3}$$

Given its monopoly power, each firm acts to maximize profit:

$$\pi_j\left(q_j\right) = p_j q_j - TC_j\left(q_j\right) \tag{4}$$

1.2 Households

Households derive utility from the consumption of goods and the services offered by land occupation (housing and amenities). These two decisions are different in nature, since, although consumption patterns can be adapted rapidly under changing economic conditions, location decisions are constrained by long-lived infrastructure and hence submitted to stronger inertias. To capture these specificities, we consider housing services and the consumption of goods as separable in the utility function, which allows us to treat them separately. We note

$$U_j(x) = U_j^{(L)}(x) + U_j^{(C)}(x) \tag{5}$$

a) Land Demand and urban costs

As traditionally approached by urban and regional economics since von Thünen (1966), land in agglomeration j is conceived as a monocentric, axisymmetric city spread along one-dimensional space $-d_j \le x \le d_j$, where d_j is the overall city size. The central business district (CBD), situated at the origin $x = 0$, is the location where firms choose to distribute once they enter the agglomeration. All economic activities take place in the j-CBD, whereas the urban population is distributed within circular peripheral areas surrounding it.

The land use component of the utility function, $U_j^{(L)}$, captures the tradeoff between the welfare gained from land consumption, assumed to be proportional to the space occupied $\lambda_j(x)$, and the amenities related to this location *A(x)*. These amenities measure the accessibility to urban services and hence decrease at higher distance x from the CBD. For the sake of simplicity, we adopt an inverse relationship, and capture substituability between land-use and amenities by a Cobb-Douglas formulation:

$$U_j^{(L)}(x) = \lambda_j(x)^{\nu_1}\left[\frac{1}{x}\right]^{\nu_2} \tag{6}$$

Households distribute in space according to an equalization of utility levels at each point x of the j-agglomeration. By introducing $\xi = \frac{\nu_2}{\nu_1}$, this means: [4]

$$\lambda_j(x) = \overline{\lambda_j} x^{\xi} \ ,0 \le \xi \le 1 \tag{7}$$

Here $\overline{\lambda_j}$ is a constant. The number of households L_j is then given by:

$$L_j = \int_{0\le|x|\le d_j} \frac{dx}{\lambda_j(x)} = \frac{2{d_j}^{1-\xi}}{\overline{\lambda_j}(1-\xi)} \tag{8}$$

Total urban costs experienced by households at a distance x from the j-CBD are made of land costs $LC_j(x) = \lambda_j(x).R_j(x)$ and commuting costs $CC_j(x)$. Commuting costs are due to the daily trip to and from the CBD. As in Murata and Thisse (2005), we introduce unitary commuting costs θ_j in the 'iceberg form' *à la* Samuelson (1952), and the effective labor supply $s_j(x)$ of a worker living in the urban area at a distance x from the CBD is[5]:

$$s_j(x) = 1 - 2\theta_j\,|x|,\ \ x \in \left[-d_j; d_j\right] \tag{9}$$

The commuting costs $CC_j(x)$ can then be expressed as the losses of revenues due to commuting: $CC_j(x) = 2\theta_j\,|x|\,w_j$. The total urban costs incurred by households for living at location x are then given by

$$UC_j(x) = 2\theta_j\,|x|\,w_j + \lambda_j(x)R_j(x) \tag{10}$$

b) Consumption of goods

Households derive utility from the consumption of the differentiated good M and the homogenous good f. Utility of a household living in agglomeration j at a distance x from the CBD is given by:

[4] Condition $\xi \ge 0$ ensures that $\lambda_j(x)$ is an increasing function, so that the empirical evidence of higher population density in the centre of the city is captured, and condition $\xi \le 1$ is necessary to have population convergence in (2).

[5] Condition: $0 < \theta_j \le \frac{1}{2d_j}$ ensures positive labor supply.

$$U_j^C(x) = U_j^{(0)} M_j(x)^{\delta_j} f_j(x)^{1-\delta_j} \tag{11}$$

Here, δ_j is the share of the manufacturing good in households' expenditures[6] , $f_j(x)$ is the consumption of the homogenous good and $M_j(x)$ is the aggregated consumption of the differentiated good, with an elasticity of substitution $\varepsilon > 1$ among the varieties. By noting $c_{kj}(x)$ the demand at a distance *x* from the *j*-CBD for a variety produced in agglomeration *k*, we have:

$$M_j(x) = \left[\sum_{k=1}^{N_A} n_k \left[c_{kj}(x) \right]^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{12}$$

The constant $U_j^{(0)}$ captures all the elements driving households' utility which are not directly related to consumption; it encompasses all amenities associated to residing in agglomeration j. In the ramining of the paper, we refer to this parameter as the "amenity for households" parameter.

By introducing the disposable income $y_j(x)$ of a household living at distance *x* from the *j*-CBD, the consumer has to satisfy the following budget constraint:

$$y_j(x) = \sum_{k=1}^{N_A} n_k p_{kj} c_{kj}(x) + p_F f_j(x) \tag{13}$$

Maximization of utility in (11) under budget constraint (13) leads to the conditions

$$i)\ f_j(x) = \left(1 - \delta_j\right) \frac{y_j(x)}{p_F}; \quad ii)\ c_{kj}(x) = \frac{\left(p_{kj}\right)^{-\varepsilon}}{I_j^{1-\varepsilon}} \left[\delta_j y_j(x) \right]. \tag{14}$$

Here we have introduced the price index of the differentiated good in agglomeration j

$$I_j = \left[\sum_{i=1}^{N_A} n_k \left(p_{kj}\right)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}. \tag{15}$$

[6] We assume that this share is agglomeration-specific, so that they may differ from one agglomeration to another but are identical for all households living in a given agglomeration.

c) Income

Households have three sources of income: wages paid to workers, dividends from capital investments and transfers from the government.
In agglomeration *j*, wages are paid on effective labor, so that a household living at a distance *x* from the CBD receives a labor-related income $y_j^{(l)}(x)$ given by:

$$y_j^{(l)}(x) = s_j(x)w_j \tag{16}$$

In agglomeration *j*, total revenues from capital $\Upsilon_j^{(K)}$ are given by: $\Upsilon_j^{(K)} = r_j X_j$. For the sake of simplicity, we assume that productive capital is equally possessed by local households, so that the dividends are uniformly re-distributed among them. Each of the L_j households living in the *j*-agglomeration receives a capital–related income $y_j^{(K)}$ given by:

$$y_j^{(K)} = \frac{r_j X_j}{L_j} \tag{17}$$

By introducing *R*$_j$*(x)* the land rent at distance *x* from the *j*-CBD, the total of land revenues perceived by governments is $\int\limits_{0 \le |x| \le d_j} R_j(x)dx$. For the sake of simplicity, we assume that the local governments redistribute this revenue in a lump-sum manner. A household living at distance *x* from the *j*-CBD pay $\lambda_j(x)R_j(x)$ and then benefits from a transfer (either positive or negative):

$$y_j^{(T)}(x) = \frac{\int\limits_{0 \le |x| \le d_j} R_j(x)dx}{L_j} - \lambda_j(x)R_j(x) \tag{18}$$

Total income in agglomeration *j* is given by

$$\Upsilon_j = \Upsilon_j^{(l)} + \Upsilon_j^{(K)} + \Upsilon_j^{(T)} \tag{19}$$

The total disposable income $y_j(x)$ of a household living at a distance *x* from the CBD is given by

$$y_j(x) = y_j^{(l)}(x) + y_j^{(K)}(x) + y_j^{(T)}(x) \tag{20}$$

1.3 Local government

Government owns the available land, and decides of the rent $R_j(x)$ to be paid by households for land use, and the amount of investments in urban infrastructure.

a) Urban Land Rent

The local government sets land rent $R_j(x)$ to ensure that people living inside each peripheral rings face identical urban costs, as in (10).[7] This means imposing

$$\forall x \in \left[-d_j ; d_j\right], \quad 2\theta_j |x| w_j + R_j(x)\lambda_j(x) = 2\theta_j d_j w_j + R_j(d_j)\lambda_j(d_j) \tag{21}$$

By normalizing the rent value at zero for land located at the edges of the city ($R_j(d_j) = 0$), the equilibrium land rent in agglomeration j is derived from equation () as:

$$R_j(x) = \frac{2\theta_j w_j \left(d_j - |x|\right)}{\lambda_j(x)} \tag{22}$$

b) Housing investments and city size

Local governments decides the amount of capital invested to construct buildings at each location x according to a minimization of total costs TC_j given by infrastructure costs IC_j and congestion costs. The latter is given by aggregate commuting costs CC_j:

$$CC_j = \int_{0 \le |x| \le d_j} \frac{CC_j(x)}{\lambda_j(x)} dx = 2\frac{1-\xi}{2-\xi}\theta_j d_j w_j L_j \tag{23}$$

Infrastructure costs IC_j are submitted to increasing marginal investment requirement at higher density in an attempt to capture higher marginal construction costs in the building sector and the need for more developed transport infrastructure. This means that the amount of annual public investment per capita I increases with density ρ :

[7] This price setting assumption ignores the monopolistic behavior of housing investors at the origin of higher rent levels. Although essential for financial flows, this dimension is less crucial for representing location patterns, which is the focus of the paper. Further extensions of the model will include this dimension through an explicit representation of housing investors.

$$I(\rho) = I_0 \rho^{\gamma-1} \tag{24}$$

Here, $\gamma > 1$ measures the non-linearity of the annual capital investments, and I_0 normalizes the units of measurment. Under condition $\gamma\xi < 1$, the total amount of annual investments in urban infrastructures in the *j*-agglomeration, IC_j, is given by:

$$IC_j = 2\int_0^{d_j} I_0 \left[\frac{1}{\lambda_j x^{\xi}}\right]^{\gamma-1} \left[\frac{1}{\lambda_j x^{\xi}}\right] dx = 2^{1-\gamma} I_0 \frac{(1-\xi)^{\gamma}}{1-\gamma\xi} L_j^{\gamma} d_j^{1-\gamma} \tag{25}$$

The minimization of total urban costs $TC_j = IC_j + CC_j$ then imposes:

$$d_j = \frac{1}{2}\left[I_0 L_j^{\gamma-1} \frac{(\gamma-1)(2-\xi)(1-\xi)^{\gamma-1}}{(1-\gamma\xi)}\right]^{\frac{1}{\gamma}} \left[\frac{1}{\theta_j w_j}\right]^{\frac{1}{\gamma}} \tag{26}$$

Equation (26) describes the combination of factors governing the dynamics of urban sprawl, beyond demography. In particular, the spatial extension d_j is inversely dependent on $\theta_j w_j$, which represents the losses of income per unit of distance traveled. This captures the incentive to adopt more dispersed settlements when the commuting distance is less penalizing, either because of lower unitary commuting costs θ_j or lower wage rate (or 'value of time') w_j. To provide an interpretation of parameter ξ, we measure the rental cost of land RC_j paid by households in agglomeration *j*:

$$RC_j = \int_{0 \le |x| \le d_j} R_j(x) dx = 2\theta_j \frac{1}{2-\xi} d_j w_j L_j \tag{27}$$

Combining (23) and (27) gives $\frac{CC_j}{RC_j} = 1-\xi$: ξ is a measure of the distribution of urban costs between commuting and housing: the lower ξ, the more commuting costs are relatively important.

2. The rural economy

2.1 Firms

In the rural area, firms produce the traditional good under constant returns to scale. Letting w_F, r_F be respectively the unitary returns of labor l_F and capital X_F, the total cost of producing q_F for a firm in the rural area is expressed as:

$$TC_F = \left[l_F w_F + r_F X_F \right] q_F \tag{28}$$

Under the perfect competition assumption, the selling price pF is set at the marginal cost of production

$$p_F = l_F w_F + r_F X_F \tag{29}$$

2.2 Households

a) Land demand

Land in the rural area is conceived as a homogenous space, where the L_A households consume the same amount of land at a price R_A.

b) Consumption of goods

The consumption part of the utility of a household living in the rural area is given by:

$$U_A^{(C)}(x) = U_A^{(0)} M_A^{\ \delta_A} f_A^{\ 1-\delta_A} \tag{30}$$

Here, $U_A^{(0)}$ is a constant, δ_A is the share of the manufacturing good in households' expenditures and f_A is the consumption of the homogenous good. Finally, M_A is the aggregated consumption of the differentiated good, with an elasticity of substitution $\varepsilon > 1$ among the varieties. By noting c_{kA} the demand in the rural area for a variety produced in agglomeration *k*, we have:

$$M_A = \left[\sum_{k=1}^{N_A} n_k \left[c_{kA} \right]^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}} \tag{31}$$

By introducing the income y_A of a household living in the rural area, the consumer has to satisfy the following budget constraint:

$$y_A = \sum_{k=1}^{N_A} n_k p_{kA} c_{kA} + p_F f_A \tag{32}$$

Maximization of utility in (30) under budget constraint (32) leads to the conditions

$$i)\ f_A = (1-\delta_A)\frac{y_A}{p_F};\quad ii)\ c_{kA} = \frac{(p_{kA})^{-\varepsilon}}{I_A^{\ 1-\varepsilon}}[\delta_A y_A]. \tag{33}$$

Here we have introduced the price index of the differentiated good in agglomeration j

$$I_A = \left[\sum_{k=1}^{N_A} n_k (p_{kA})^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}. \tag{34}$$

c) Income

Households have three sources of income: wages paid to workers, dividends from capital investments and transfers from the government.

Total labor-related and capital-related incomes are given respectively by:

$$\begin{cases} \Upsilon_A^{(l)} = w_F l_F q_F \\ \Upsilon_A^{(K)} = r_F X_F q_F \end{cases} \tag{35}$$

Finally, all rural households consume the same amount of land at the same price, and these land costs are fully compensated by equal redistribution of this rent value by the government. This means $y_A^{(T)} = 0$

Since households in the rural area are identical, each household receives an income y_A given by:

$$y_A = \frac{\Upsilon_A^{(l)} + \Upsilon_A^{(K)}}{L_A} \tag{36}$$

3. Interregional Trade

In order to allow the model for the spatial dimension, trade is allowed across agglomerations, as well as between agglomerations and the rural area. We use the 'iceberg' form of transport costs associated with trade of the composite goods (Samuelson, 1952). In particular, if one variety *i* of manufactured goods is shipped from agglomeration *j* to agglomeration *k* (to the rural area), only a fraction τ_{jk} (τ_{jA}) will arrive at the destination, the remainder melting during the shipment. To ensure that any unit produced in agglomeration *j* provides the same revenue independently from the location where it is sold, a variety sold at price p_j in its production location *j* will be charged in consumption location *k* at a price p_{jk} given by:

$$p_{jk} = \tau_{jk} p_j \tag{37}$$

Similarly, a variety produced in agglomeration *j* and sold in the rural area will be charged at a price p_{jA} given by:

$$p_{jA} = \tau_{jA} p_j \tag{38}$$

We assume that this 'traditional' good is freely traded across regions, so that its selling price in agglomeration *j*, p_{Fj}, is identical in all agglomerations and equals the selling price in the rural area p_F where it is produced: $p_{Fj} = p_F$.

4. Equilibrium

4.1 Market equilibrium for the differentiated good

The production size q_j of a firm located in agglomeration *j* must equal the sum of local consumption and exports towards other regions of the variety it produces. The market clearance condition then imposes:

$$q_j = \sum_{k=1}^{N_A} \int_{x=-d_k}^{x=d_k} \tau_{jk} c_{jk}(x) \frac{dx}{\lambda_k(x)} + \tau_{jA} c_{jA} L_A \tag{39}$$

The first term and second terms on the right-hand side are the volume of goods exported towards agglomerations and the rural area, respectively, including the amount that melts during the shipment.[8]

Under Dixit-Stiglitz monopolistic market, firms set their price by assuming a constant elasticity of substitution (CES), $\varepsilon > 1$, and profit maximization leads to a constant mark-up on variable cost: $p_j = \frac{\varepsilon}{\varepsilon - 1} \frac{\partial TC_j}{\partial q_j}$. With (3), this leads to:

$$p_j = \frac{\varepsilon}{\varepsilon - 1} \frac{l_j w_j}{n_j^{\alpha}} \tag{40}$$

As a consequence of the profit maximization behavior, the number of firms in agglomeration j is such that profits are zero, as an equilibrium condition of monopolistic competition. Therefore, by setting zero profit in (4), the return to capital r_j at the equilibrium is:

$$r_j = \frac{q_j}{\chi} \left[p_j - \frac{l_j w_j}{n_j^{\alpha}} \right] \tag{41}$$

4.2 Market equilibrium for the homogenous good

For the homogenous good produced in the rural area, market clearing imposes that:

$$q_F = \sum_{k=1}^{N_A} \int_{x=-d_j/2}^{x=d_j/2} f_k(x) \frac{dx}{\lambda_k(x)} + L_A f_A \tag{42}$$

Here, the first term on the right-hand side is the consumption of the traditional good from households residing in agglomeration j, whereas the second term represents total consumption from households of the rural area.

The assumption of perfect competition imposes marginal cost pricing: $p_F = \frac{\partial TC_F}{\partial q_F}$. With (28) this leads to:

[8] We adopt the natural convention that $\tau_{jj} = 1$ (no trade cost for a good produced and consumed in the same region)

$$p_F = l_F w_F + r_F X_F \tag{43}$$

4.3 Labor market equilibrium

Total effective labor supply S_j in the j-agglomeration is given by:

$$S_j = \int_{0\le|x|\le d_j} \frac{s_j(x)dx}{\lambda_j(x)} = L_j\left[1-2\frac{1-\xi}{2-\xi}\theta_j d_j\right] \tag{44}$$

Total labor requirement for production in agglomeration j is given by $\frac{l_j}{n_j^{\alpha}} n_j q_j$. At the labor-market equilibrium, it must equal total labor supply in this agglomeration j:

$$S_j = \frac{l_j}{n_j^{\alpha}} n_j q_j \tag{45}$$

Total labor requirement for production of the homogenous good is given by $l_F q_F$. At the labor-market equilibrium, it must equal the total labor effectively supplied L_A :

$$L_A = l_F q_F \tag{46}$$

4.4 Land market equilibrium across regions

Combining (11) and (14), the utility level $\tilde{U}_j$ in the j-agglomeration is equal to

$$\tilde{U}_j = U_j^{(0)} \frac{\left[\delta_j^{\delta_j}\left(1-\delta_j\right)^{1-\delta_j}\right]}{I_j^{\delta_j} p_F^{1-\delta_j}} \frac{p_j n_j q_j}{L_j} \tag{47}$$

Similarly, the utility level $\tilde{U}_A$ in the rural area is given by

$$\tilde{U}_A = U_A^{(0)} \frac{\left[\delta_A^{\delta_A}\left(1-\delta_A\right)^{1-\delta_A}\right]}{I_A^{\delta_A} p_F^{1-\delta_A}} \frac{p_F q_F}{L_A} \tag{48}$$

As it is standard in New Economic Geography models *à la* Krugman, workers have an incentive to move towards locations where they get the higher utility and migration decisions are based on utility differentials across different regions. At the equilibrium, workers have no incentive to relocate, which imposes that they reach the same utility level in any of the N_A+1 regions. This means imposing : $\forall j \in [1 \;\; N_A], \quad \tilde{U}_j = \tilde{U}_A$. With (47) and (48), this leads to :

$$\forall j \in [1 \;\; N_A], \quad U_j^{(0)} \frac{\left[\delta_j^{\delta_j}\left(1-\delta_j\right)^{1-\delta_j}\right]}{I_j^{\delta_j} p_F^{1-\delta_j}} \frac{p_j n_j q_j}{L_j} = U_A^{(0)} \frac{\left[\delta_A^{\delta_A}\left(1-\delta_A\right)^{1-\delta_A}\right]}{I_A^{\delta_A} p_F^{1-\delta_A}} \frac{p_F q_F}{L_A} \tag{49}$$

This concludes the short-run model.

II- Dynamics

The set of equations (1)-(49) describes the equilibrium of the economy at a given date t[9]. We adopt a recursive structure in which the determinants of spatial locations are modified between dates *t* and *t+1* by aggregate macroeconomic trends at the national/international level and firms' location choices among the cities in interaction.

1- Macroeconomic trends

The spatial model is conceived to spatially disaggregate the economy of a (group of) region(s). The disaggregate economies at the urban scale and the aggregate macroeconomy at the country scale are consistent if each variable appearing in both the scale description are such that the aggregation of the former equals the latter. This imposes a set of consistency equations to ensure that, at each date *t*, the spatial economy, partitioned into a number of agglomerations and a rural area, is consistent with major aggregate macroeconomic trends. We introduce $\overline{V(t)}$, $\overline{S(t)}$, $\overline{w(t)}$, $\overline{Pop(t)}$ the total value of production, total working force (effective labor), average wage rate and national population, respectively. Consistency between the two descriptions imposes a set of equations at each date *t*:

$$\sum_{k=1}^{N_A} p_k n_k q_k + p_F q_F = \overline{V(t)} \tag{50}$$

$$\sum_{k=1}^{N_A} S_k + l_F q_F = \overline{S(t)} \tag{51}$$

$$\sum_{k=1}^{N_A} w_k S_k + w_F l_F q_F = \overline{w(t)}\,\overline{S(t)} \tag{52}$$

[9] In the reminder of the paper, the time-dependence of spatial variables will be kept implicit to simplify the notations

$$\sum_{k=1}^{N_A} Pop_k + Pop_A = \overline{Pop(t)} \tag{53}$$

2- Firms' location choices

The second dynamic mechanism comes from firms' location decisions and induced changes in the spatial distribution of production and productive capital in the national economy. Agglomerations differ in labor force, infrastructure endowment and amenities as captured by labor productivity l_j, unitary commuting costs θ_j, and utility factor $U_j^{(0)}$, respectively. These *j*-specificities act as constraints on the local economies, and influence in particular the attractiveness of the *j*-agglomeration for productive capital. As a measure of investment profitability, this attractiveness A_j is expected to be principally dependent upon the return to capital r_j. Section IV tests this hypothesis on past historical data.

The agglomeration attractiveness $A_j(t)$ acts as a driving force of firms' migration decisions, which have an incentive to settle in agglomerations with higher attractiveness. This dynamic effect is captured by assuming that the relative variation of the number of firms in a given agglomeration is an increasing function of its attractiveness index. Assuming linear dependence for the sake of simplicity, we have:

$$\frac{\Delta n_j(t)}{n_j(t)} = Max\left(A_j(t) - \overline{A}, -1 \right) \tag{54}$$

The condition $\frac{\Delta n_j(t)}{n_j(t)} > -1$ ensures positive number of firms, and $\overline{A}$ is a parameter controlling the total variation of the number of firms Δn:

$$\overline{A} = \frac{\sum_j A_j n_j - \Delta n}{\sum_j n_j} \tag{55}$$

The number of firms in agglomeration *j* at date *t+1* is then given by:

$$n_j(t+1) = n_j(t) + \Delta n_j(t) \tag{56}$$

III- Base-year Calibration

The purpose of carrying out the base-year calibration is to provide all the variables defining the model with numerical values that enables the economy to be representative of the reality at the baseline year (2001 in our case), both at the national/regional level and at the urban scale.

The national/regional dimension is operationalized on OECD regions divided into four macro-regions: USA, Canada, Europe[10] (EUR) and OECD Pacific[11]. In each of them, the consistency of the disaggregated urban economies with global macroeconomic flows at the base year impose the following equations on the disaggregate variables at the base year

$$\sum_{k=1}^{N_A} p_k n_k q_k + p_F q_F = \overline{V^{(0)}} \tag{57}$$

$$\sum_{k=1}^{N_A} S_k + l_F q_F = \overline{S^{(0)}} \tag{58}$$

$$\sum_{k=1}^{N_A} w_k S_k + w_F l_F q_F = \overline{w^{(0)}}\,\overline{S^{(0)}} \tag{59}$$

$$\sum_{k=1}^{N_A} Pop_k + Pop_A = \overline{Pop^{(0)}} \tag{60}$$

This set of equation is similar to (50)-(53), but is applied at the base year when the total value of production, the total working force (effective labor), the average wage rate and the national population take their values at date 2001 (t=0), $\overline{V^{(0)}}$, $\overline{S^{(0)}}$, $\overline{w^{(0)}}$ and $\overline{Pop^{(0)}}$, respectively. The macroeconomic aggregates $\overline{V^{(0)}}$, $\overline{S^{(0)}}$ and $\overline{w^{(0)}}$ are derived from the GTAP-6 database, which provides, for the year 2001, a set of balanced input-output tables of the world economy, detailed in 87 regions and 57 sectors (Dimaranan et al., 2006). Demographic data for total population $\overline{Pop^{(0)}}$ are taken from (UN, 2005) (Table C-2 in Appendix).

Equations (57)-(60) control the systems of cities in interaction as a whole, but are insufficient to provide a detailed differentiation of economic activity among the different agglomerations

[10] EU-27+ Turkey
[11] Japan, South Korea, Australia and New Zealand

that compose this urban system. To do so, the standard calibration process at the macroeconomic level is complemented by a calibration of the agglomeration-specific variables. This second stage of calibration consists in imposing that, for each agglomeration j, the crucial socio-economic variables at the agglomeration level are consistent with their values observed in the reality. We retain five local characteristics: the number of households $L_j^{(0)}$, the spatial extension $d_j^{(0)}$, the wage rate $w_j^{(0)}$, the production $Q_j^{(0)}$ and the commuting cost $CC_j^{(0)}$. For any j-agglomeration of the urban panel under consideration, we impose the following fitting equations on local data:

$$L_j = L_j^{(0)} \tag{61}$$

$$d_j = d_j^{(0)} \tag{62}$$

$$\frac{w_j}{w_1} = \frac{w_j^{(0)}}{w_1^{(0)}} \tag{63}$$

$$\frac{n_j q_j}{n_1 q_1} = \frac{Q_j^{(0)}}{Q_1^{(0)}} \tag{64}$$

$$\frac{S_j}{L_j} = CC_j^{(0)} \tag{65}$$

In this study, the agglomerations considered are taken from the Metropolitan Database of the OECD and include 74 among the most important OECD agglomerations, representing 37% of OECD population and 48% of OECD GDP (the complete list of those agglomerations is given in Appendix A). In equations (61)-(65), the number of households $L_j^{(0)}$, the spatial extension $d_j^{(0)}$ and the relative production levels $Q_j^{(0)}$ are given by the Metropolitan Database of the OECD, in accordance with their definition of the metro-regions[12], and relative wages $w_j^{(0)}$ are derived from a study by UBS[13] (see Table C-1 in the Appendix).

[12]available at : http://stats.oecd.org/Index.aspx?DataSetCode=METRO
[13] available at : *http://www.ubs.com/1/e/wealthmanagement/wealth_management_research/prices_earnings.html*

IV- Dynamic Calibration

We test the robustness of the model against historical data on development patterns at the local/urban scale for the 74 metropolitan regions under consideration. To this aim, starting from its calibration date 2001 (see section III), the model is run backwards over the period 1980-2001 by following its endogenous dynamics for negative times. We represent respectively the macroeconomic trends and urban dynamics over the period by imposing at each date from 2001 to 1980:

- the macroeconomic aggregates $M \in \{V, l, w, Pop\}$ in equations (50)-(53), as recursively defined for negative times by the generic equation $\overline{M(t-1)} = \frac{\overline{M(t)}}{1+g_M(t)}$, with g_M giving the average growth rate (Table A-3 in Appendix).
- past demographic trends at the urban level, as derived from city-level population growth rates in the WUP Database of the UN (UN, 2007) (see Figure C-1 in Appendix)

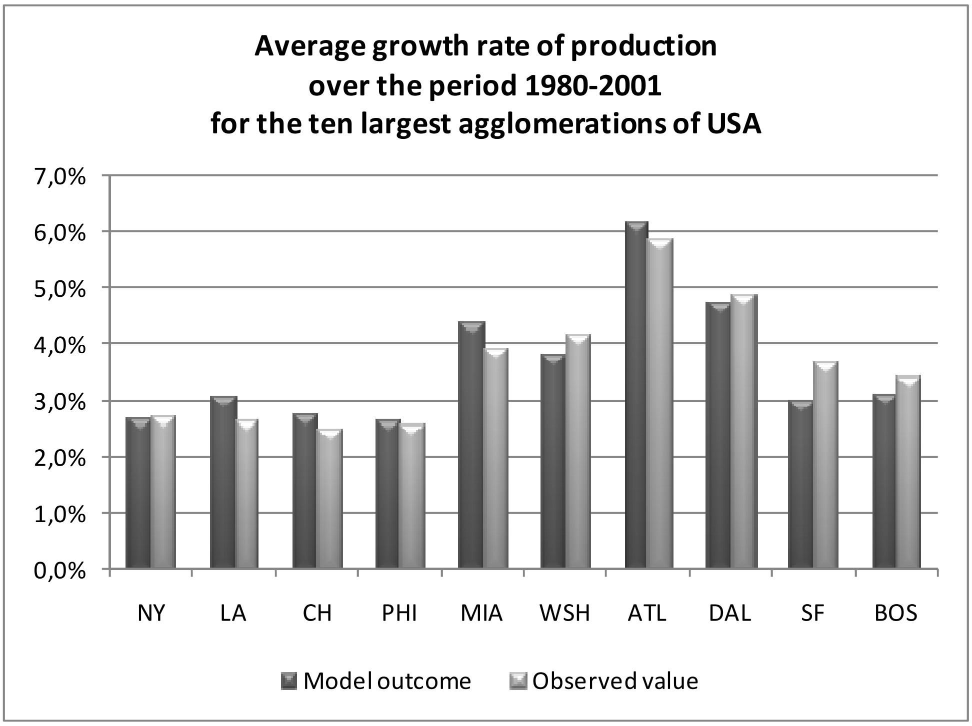
This analysis has the twofold objective of *(i)* validating the functioning of the model by demonstrating that its endogenous dynamics succeeds in reproducing past trends, and *(ii)* revealing the missing parameters of the attractiveness index that are consistent with past observed dynamics. This second step serves as a dynamic calibration procedure, since it helps to give value to parameters driving the dynamics.

1. Validation of the functioning of the model on past trends

The endogenous functioning of the model in this exercise on the 1980-2001 period gives the characteristics of local economies that are consistent with observed macroeconomic and population trends. Here, we compare the result on local production as predicted by the model with observed values over the period. [14] For the sake of simplicity, we limit our analysis to the ten largest US agglomerations (Figure 5.1)

[14] Data for real production are derived from GDP statistics from the Bureau of Economic Analysis (http://www.bea.gov/regional/reis/default.cfm?selTable=CA1-3&section=2) and price index statistics from the Bureau of Labor Statistics (http://data.bls.gov/cgi-bin/surveymost?cu).

Figure 5.1: *Average growth rate of production for the ten largest agglomerations in USA*



The endogenous relocation of production triggered by firms' migration decisions appears to be highly consistent with empirical data for real production growth available for US agglomerations. This result reinforces the credibility of model outcomes for long-term forecasts, since it demonstrates that the functioning of the model ensures a consistent description of the interactions between demographic and economic dimensions of the urban agglomerations and the spatial organization of economic activities at the national level.

The gaps between model outcomes and observed values for some agglomerations (e.g. Los Angeles, San Francisco) are due to the assumption that macroeconomic productivity gains are homogeneously distributed among agglomerations and hence that productivity differentials remain unchanged during the dynamics. This ignores the agglomeration-specific drivers of labor productivity gains, which may trigger distortions of labor productivity levels and affect migration decisions. Going further into this dimension would require entering into the details of productivity drivers in each agglomeration, which is far beyond the scope of this paper.

2. Revealing 'missing' parameters driving the dynamics

The observation of past trends helps to reveal the determinants of urban dynamics and in particular of firms' relocation decisions. Indeed, the population and production migrations are triggered by different levels of attractiveness among agglomerations, consistently with equation (54). The analysis of the attractiveness index during the period 1980-2001 helps identify its underlying determinants, and, in particular, test the assumption that it is mainly determined by the return to capital r_j . To this aim, we introduce $A_j^{(0)}$ the residual of attractiveness discounted from the return to capital, as follows:

$$A_j^{(0)}(t) = A_j(t) - r_j(t) \tag{66}$$

In equation (66), the return to capital r_j captures the incentive for investors to move to a location where profit expectations are higher, whereras parameter $A_j^{(0)}$ captures the other *j*-specificities that may either attract or discourage investments (infrastructure, local governance, nature of the working force…). We refer to this latter as the "external advantages for investors" parameter and measure its relative importance in driving attractiveness through the ratio $\left|\frac{A_j^{(0)}}{r_j}\right|$: a value close to zero means that the return to capital is the major determinant of attractiveness indexes, whereas higher-than-one values correspond to a dominant importance of other factors. This ratio is calculated in equation (66) at each date of the 1980-2001 period from the endogenous return to capital (equation 41) and the attractiveness index corresponding to endogenous firms' migrations (equation 54), both being consistent with macroeconomic and urban population forcings.

The results prove that the ration remains significantly below one, with a maximum value of 0.23 and an average of 0.03 over all agglomerations and all dates between 1980 and 2001. This means that the return to capital explains on average 97% of the agglomeration attractiveness, and justifies the focus on this determinant when describing the attractiveness. Parameter $A_j^{(0)}$ takes differentiated values across agglomerations with weak magnitude of time variations and Table 5.1 reports the average value of the $A_j^{(0)}$ parameter over 1980-2001 for each agglomerations.

Table 5.1: *Average values of* $A_j^{(0)}$ *for the largest agglomerations, over the period 1980-2001*

USA	NY	LA	CH	PHI	MIA	WSH	ATL	DAL	SF	BOS
	0.037	0.024	0.006	0.004	0.004	-0.014	-0.026	-0.04	0.054	-0.012
Canada	TOR	MONT	VANC							
	0.001	0.005	-0.005							
Europe	RH	IST	PAR	RAND	MIL	LON	MUN	BER	FKT	MAD
	0.009	-0.016	0.029	0.003	-0.065	0.01	0.021	0.01	0.011	0.008
OECD Pacific	TOK	SEO	OSA	AIC	BUS	FUK	SYD	MEL	DEA	AUC
	0.025	-0.059	0.030	0.024	-0.038	0.018	0.032	0.033	-0.059	-0.027

In the remaining of the paper, we consider the above values of $A_j^{(0)}$ as a constant-over-time additive component in the attractiveness index of each agglomeration. This parameter helps capture the effects on long-term dynamics of some dimensions that are not explicitly represented in the model but affect migration decisions in long-term forecasts. In this sense, it constitutes a dynamic calibration of the model in that it derives the value of a crucial parameter for the dynamics from the comparaison between the endogenous properties of the model and dynamic observations.

V-Long-term equilibrium

To characterize the model mechanisms in the long-term, we study the long-term equilibrium of the model resulting from firms' migration decisions, everything else being kept constant. Here, we consider that both macroeconomic aggregates $M \in \{V, l, w, X, L\}$ and agglomeration-specific parameters are constant and we analyze the number of firms in each agglomeration, which progressively evolves from its initial value to a long-term agglomeration-specific level after arbitrary long time period[15]. This latter situation in which agents have no incentive to migrate defines the long-term equilibrium.

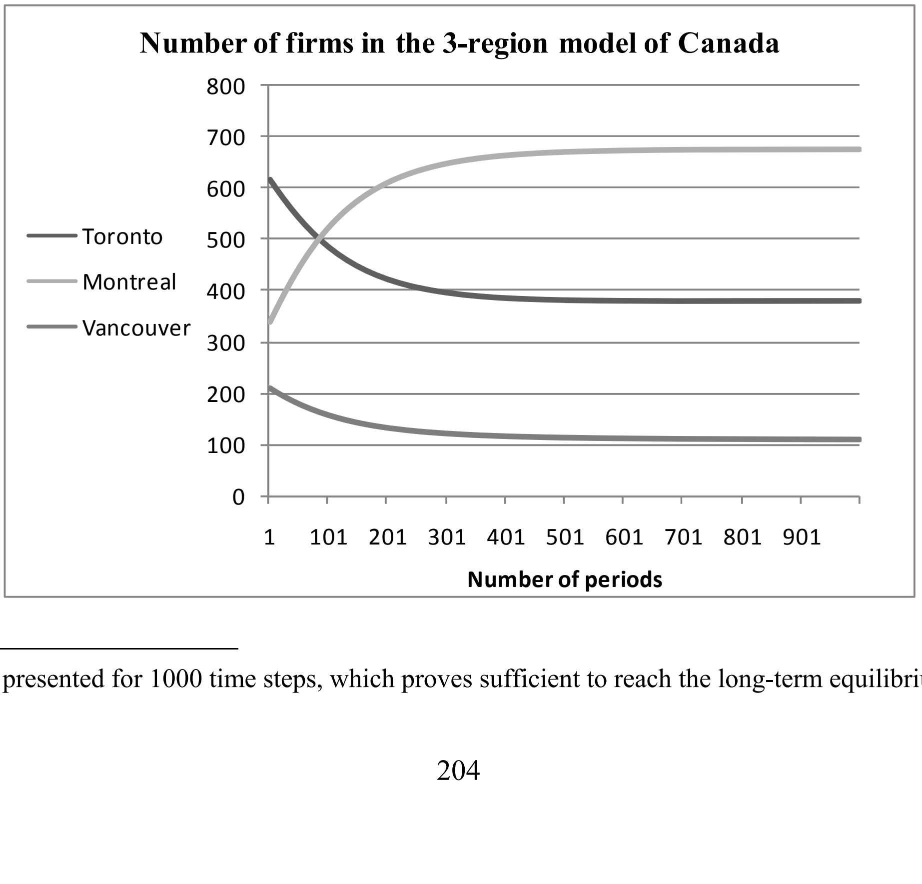
In this section, we analyze the dependency of these long-term equilibria to trade barriers. For the sake of simplicity, we assume homogenous for (i.e., τ_{jk} is identical for any pair of regions, $j \neq k$) and we note τ this value. As it is standard in NEG analyses, we start by considering the two extreme cases, free trade ($\tau = 1$) and autarky ($\tau \to \infty$), before deriving the 'bifurcation diagram' over the whole range of trade barriers.

1. Long-term equilibrium under free trade ($\tau = 1$)

1.1 Numerical results

We assume free trade and, starting from the calibration date 2001, we consider the dynamics driven by firms' migration decisions towards the most attractive locations. Figure 5.2 demonstrates 'partial agglomeration' equilibrium for Canadian agglomerations, with a long-run distribution of firms across several agglomerations. Figures C-2, C-3 and C-4 in Appendix C illustrate the same outcome under the free trade assumption for the USA, Europe and OECD Pacific agglomerations, respectively. This is consistent with findings from Chapter 4 which demonstrates that partial equilibria with dispersion of production at free trade are possible when considering the heterogeneity of intra-regional spatial organization.

Figure 5.2: *Dynamic trends of the three-region model of Canada over 1000 periods for* $\tau = 1$



[15] Results are presented for 1000 time steps, which proves sufficient to reach the long-term equilibrium for $\tau = 1$.

Table 5.2 summarizes the numerical results under the free trade assumption by giving the spatial distribution of firms in the long-run equilibrium for the four macro-regions.

Table 5.2: *Share of the number of firms in the long-run equilibrium for* $\tau = 1$ *(%)*

USA*	NY	LA	CH	PHI	MIA	WSH	ATL	DAL	SF	BOS
	0.17	0.11	0.13	0.08	0.00	0.05	0.00	0.04	0.00	0.07
Canada	TOR	MONT	VANC							
	0.33	0.58	0.09							
Europe*	RH	IST	PAR	RAND	MIL	LON	MUN	BER	FKT	MAD
	0.23	0.00	0.08	0.00	0.12	0.07	0.08	0.01	0.05	0.00
OECD Pacific	TOK	SEO	OSA	AIC	BUS	FUK	SYD	MEL	DEA	AUC
	0.34	0.25	0.22	0.08	0.10	0.00	0.00	0.00	0.01	0.00

**Note: the total does not sum up to 1 since only the results from the ten largest agglomerations (out of 23 and 38 for USA and Europe respectively) are reported.*

1.2 Analytical solution

To analyze more in-depth the determinants of these results, we provide an analytical solution of the long-run model in the specific case where $\alpha = 0$ (negligible agglomeration effect) and income shares on manufacturing goods are identical in all regions (we note δ this value). These assumptions are adopted to ensure analytical tractability of the model, but do not undermine the explanatory power of the analysis, which aims at making explicit the determinants of the number of firms at the long-term equilibrium.

The free trade assumption implies that the price index is identical in all regions (agglomerations and rural area):

$$I = \left[\sum_{k=1}^{N_A} n_k \left(p_k \right)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} \tag{67}$$

By introducing Υ is the total households' income, a direct calculation with equation (39) and (67) gives the production level:

$$q_j = \frac{p_j^{-\varepsilon}}{I^{1-\varepsilon}} \delta \Upsilon \tag{68}$$

From (40), (41) and (68), we derive the return to capital:

$$r_j = \frac{\delta \Upsilon}{\varepsilon \chi I^{1-\varepsilon}} p_j^{1-\varepsilon} . \tag{69}$$

By introducing $K_1 = \frac{U_A^{(0)}}{\varepsilon \chi L_A} (1-\delta) \Upsilon$, equation (49) defining equilibrium among different locations can then be rewritten as:

$$\frac{n_j}{L_j} = \frac{K_1}{U_j^{(0)} r_j} \tag{70}$$

Combining equations (26), (40), (46) and (68) and introducing the constant $K_2 = \left[I_0 \frac{\varepsilon}{\varepsilon - 1} \left(\frac{1-\xi}{2-\xi} \right)^{\gamma} \frac{(\gamma-1)(2-\xi)(1-\xi)^{\gamma-1}}{(1-\gamma\xi)} \right]^{\frac{1}{\gamma}}$, the equilibrium on urban labor markets imposes condition:

$$L_j \left[1 - K_2 \left(\theta_j\right)^{\frac{\gamma-1}{\gamma}} \left(L_j\right)^{\frac{\gamma-1}{\gamma}} \left(l_j\right)^{\frac{1}{\gamma}} \left(p_j\right)^{-\frac{1}{\gamma}} \right] = \frac{\delta \Upsilon}{I^{1-\varepsilon}} l_j n_j \left(p_j\right)^{-\varepsilon} \tag{71}$$

Combining (69), (70) and (71), and introducing the two parameters $K_3 = K_1 \left(\frac{I^{1-\varepsilon}}{\delta \Upsilon} \right)^{\frac{1}{\varepsilon-1}} \left(\varepsilon\chi\right)^{\frac{\varepsilon}{\varepsilon-1}}$ and $K_4 = K_2 \left(\frac{\varepsilon \chi I^{1-\varepsilon}}{\delta \Upsilon} \right)^{\frac{1}{\gamma(\varepsilon-1)}}$, we obtain:

$$\left(L_j\right)^{\frac{\gamma-1}{\gamma}} = \frac{1 - K_3 \frac{l_j \left(r_j\right)^{\frac{1}{\varepsilon-1}}}{U_j^{(0)}}}{K_4 \left(\theta_j\right)^{\frac{\gamma-1}{\gamma}} \left(l_j\right)^{\frac{1}{\gamma}} \left(r_j\right)^{\frac{1}{\gamma(\varepsilon-1)}}} \tag{72}$$

The long-term partial equilibrium is defined by the equality of all attractiveness indexes at a common value $\overline{A}$. With (66), this means that the return to capital r_j at the long-term partial equilibrium satisfy

$$r_j = \overline{A} - A_j^{(0)}. \tag{73}$$

Plugging (73) into (72) gives the population in agglomeration j at the long-term partial equilibrium in function of the parameters characterizing the j-agglomeration:

$$L_j = \frac{\left[1 - K_3 \dfrac{l_j}{U_j^{(0)}} \left(\overline{A} - A_j^{(0)}\right)^{\frac{1}{\varepsilon-1}}\right]^{\frac{\gamma}{\gamma-1}}}{K_4 \theta_j \left(l_j\right)^{\frac{1}{\gamma-1}} \left(\overline{A} - A_j^{(0)}\right)^{\frac{1}{(\gamma-1)(\varepsilon-1)}}} \tag{74}$$

This equation represents the analytical solution of the model by providing an explicit relation between the population in agglomeration j at the long-term equilibrium and the parameters characteristizing this agglomeration, which are kept constant at their base-year value. A direct analysis of equation (74) demonstrates that the long-run population is higher in agglomerations where, everything else being equal, labor productivity is higher (lower l_j), unitary commuting costs are lower (lower θ_j), amenities for households are higher (higher $U_j^{(0)}$) or external advantages for investors are higher (higher $A_j^{(0)}$). The long-run outcomes in Table 2 result from the tradeoff between these four agglomeration-specific characteristics, as described in (74).

2. Long-term equilibrium under autarky ($\tau \to \infty$)

In autarky, the price index in agglomeration j and in the rural area, I_j and I_A, can be approximated by:

$$(a)\; I_j^{1-\varepsilon} = n_j p_j^{1-\varepsilon}; \quad (b)\; I_A^{1-\varepsilon} = \tau^{1-\varepsilon} \sum_{k=1}^{N_A} n_k p_k^{1-\varepsilon} \tag{75}$$

Equation (39) giving the production level q_j and equation (41) giving the return to capital r_j in agglomeration j can be rewritten as:

$$(a)\, q_j = \frac{p_j^{-\varepsilon}}{\sum_k n_k p_k^{1-\varepsilon}} \delta \Upsilon; \quad (b)\, r_j = \frac{1}{\varepsilon \chi} \frac{p_j^{1-\varepsilon}}{\sum_k n_k p_k^{1-\varepsilon}} \tag{76}$$

From (75) and (76), the equilibrium condition (49) between the rural area and a given agglomeration j gives:

$$\frac{U_j^{(0)}}{L_j} \left[n_j r_j \right]^{1+\frac{\delta}{\varepsilon-1}} = \frac{1}{\tau^{\delta}} \frac{1-\delta}{\delta \left[\varepsilon \chi \right]^{1+\frac{\delta}{\varepsilon-1}}} \frac{U_A^{(0)}}{L_A} \tag{77}$$

Let us assume that a partial equilibrium exists, so that at least two different agglomerations $j_1 \neq j_2$ (with $A_{j_1}^{(0)} \neq A_{j_2}^{(0)}$) are non-empty in the long-term ($n_{j_1} \neq 0$ and $n_{j_2} \neq 0$).

- Equation (77) then imposes $r_{j_1} = r_{j_2} = 0$ (*)
- The equilibrium condition in the long-term imposes $A_{j_1} = A_{j_2}$. With (66), this means $r_{j_1} + A_{j_1}^{(0)} = r_{j_2} + A_{j_2}^{(0)}$ (**)
- Combining (*) and (**), we have $A_{j_1}^{(0)} = A_{j_2}^{(0)}$, which is in contradiction with the initial condition. This means that the assumption of a partial equilibrium is false and that the long-term equilibrium is necessarily a full agglomeration in a unique agglomeration.

This result is opposite to the standard finding of the NEG literature that autarky corresponds to a symmetric distribution of production among the regions (Krugman, 1991). This difference is caused by a different representation of the location of production. In standard NEG approaches, all the regions considered produce both types of goods, so that local production can provide all the types of goods demanded by local households. At high trade costs, households have no incentive to demand goods produced in the other region and consume only the varieties of the manufacturing good that are produced locally. In our model on the contrary, we take into account the differentiation of production according to the region considered and manufacturing goods can be produced only in agglomerations. This means in particular that, even at high trade costs, it remains necessary to trade manufacturing goods produced in agglomerations in order to satisfy households' demand in the rural area. This necessity to consume traded goods significantly increases the price index in the rural area in case of high trade costs (see equation 75(b)), and the concentration of all manufacturing production in a single agglomeration is then a way to limit its production cost and then the rural price index.

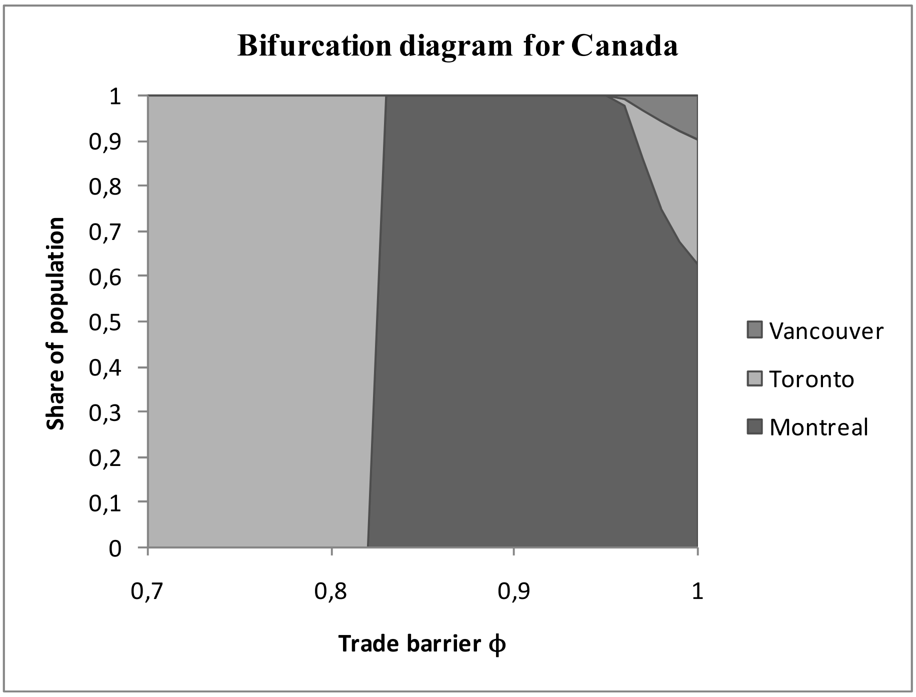
3. The bifurcation diagram

We now investigate more generally the long-term equilibria over the whole range of trade freeness $\phi = \tau^{1-\varepsilon}$ ($0 \le \phi \le 1$), and summarize the results in a 'bifurcation diagram'.[16] Such representation is conventionnally used to summarize the outcomes of 2-region NEG models, and the present analysis extends this approach to the context of multiple regions (or agglomerations)

The 'bifurcation diagram' for USA, Europe and OECD Pacific are provided in Appendix (Figures C-5, C-6 and C-7, respectively). These results demonstrate two common general characteristics. On the one hand, there exists a range of trade barriers (depending on the region considered) which is consistent with partial agglomeration favoring a distribution of economic activity over many of the agglomerations. On the other hand, the decrease of trade barrier ϕ enhances the agglomeration process by favoring concentration into a more and more reduced number of agglomerations; this is consistent with the previous analysis, which demonstrates full agglomeration in the extreme case of autarky ϕ=0. Due to the complexity of the interactions in these three regions (which include 23, 38 and 10 agglomeration in interaction for USA, Europe and OECD Pacific, respectively), we do not enter further into analytical investigation of the determinants of these results. We do so in the more simple case of Canada, which has the advantage of including a reduced number of agglomerations making it possible to disentangle the mechanisms at play. Figure 5.3 gives the 'bifurcation diagram' in this context.

[16] The numerical results are presented over the range $0.7 \le \phi \le 1$ for the sake of readibility of the graphics. This range proves sufficient to illustrate the more important trends demonstrated by the analysis

Figure 5.3: *Bifurcation diagram for Canada*



There exists a range of trade barriers over which we obtain a partial long-run equilibrium, where firms distribute over several agglomerations. The existence of such 'partial equilibrium' demonstrates a step-forward towards the empirical validity of NEG models, by making it possible that the multiple agglomeration interactions gives rise to an equilibrium in which production is distributed among a number of agglomerations, in line with empirical observations. The consistency of this picture with empirical observations is reinforced by the fact that these partial equilibria are obtained for sufficiently high trade barriers, corresponding to low trade costs ($\phi \geq 0.95$ or $\tau \leq 1.22$), in line with observed patterns of low obstacles to trade in the economy. To understand the respective shares of the three agglomerations in the long-run distribution, we go back to the inherent characteristics of the agglomerations, which remain constant-in-time in this analysis of long-run equilibria and determine firms' location decisions: the unitary labor requirements for production l_j (inverse of labor productivity), the unitary commuting costs θ_j, the amenity for households $U_j^{(0)}$ and the external advantages for investors $A_j^{(0)}$ (Table 5.3).

Vancouver is disfavored with respect to the other agglomeration for all the criteria, which explains why it aways take the smaller share of population and economic activity. Toronto performs best in terms of production costs by offering higher productivity (lower l_j) and lower commuting costs (lower θ_j), but Montreal is favored in terms of implicit determinants, both for households' amenities (higher $U_j^{(0)}$) and external advantages for investors (higher $A_j^{(0)}$). Toronto has lower costs ensuring a competitive advantage over varieties produced in other agglomerations, higher volumes of sales and higher return-to-capital; but, the low trade cost assumption limits the multiplying effect on the final price of goods and hence the gap in relative prices (see equation (37)) and the difference in return-to-capital. In this case, the exogenous parameters (which are in favor of Montreal) play a crucial role by more than offsetting the above endogenous effects in the attractiveness to foster migration towards Montreal, as demonstrated by the dominant share it takes for $\phi \geq 0.95$.

Table 5.3: *Characteristics of the Canadian agglomerations (arbitrary units)*

	Unitary labor requirements for production l_j	Unitary commuting costs θ_j	Amenity for households $U_j^{(0)}$	External advantages for investors $A_j^{(0)}$
Montreal	9.7	5.2	**0.86**	**4.6**
Toronto	**8.4**	**4.4**	0.70	0.5
Vancouver	9.4	6.4	0.84	-5.1

For lower trade barriers $\phi \leq 0.95$ representing high trade costs, a full agglomeration pattern is obtained. Interestingly, the 3-region case of Canada demonstrates that the agglomeration which attracts all firms may depend on the level of trade freeness. For ϕ-values (still lower but) close to 0.95, the tradeoff between determinants of migration are similarly in favor of Montreal like in the partial agglomeration case described above; the only difference is that higher trade costs foster full concentration of production allowing to satisfy urban consumption through local production, and hence avoiding losses due to inter-agglomerations trade. But, at sufficiently high trade costs, firms prove to concentrate in Toronto. This change in concentration patterns is an original outcome of our exercise and is related to the trade patterns between the agglomeration and the rural area. Indeed, the differentiated good is produced only in the agglomerations and part of them must be traded from agglomerations to the rural area to satisfy rural households' demand even at high trade costs (low ϕ-values). When the cost of trading urban goods towards the rural area is very high, producers have interest to reduce production costs to stabilize rural consumers' price index and hence maintain the amount of goods sold in the rural area and the associated benefits. This means favoring the agglomeration where production costs (and hence selling prices) are lowest, Toronto in our case.

VI-Long-term forecasts

Let us now turn to consider the effect of macroeconomic dynamics on urban development by solving recursively the urban model given by equations (1)-(49) on a yearly basis. This means that, at each date of the period 2001-2050, the macroeconomic aggregate variables $M \in \{V, S, w, Pop\}$ in consistency equations (50)-(53) are given by exogenous trends (see Table 5.4 for values in intermediate years) and that we investigate the effect of these macroeconomic variables on endogenous firms' location decisions triggered by differences in attractiveness among agglomerations given by (54).

We adopt three simplifying assumptions. First, we assume constant values of agglomeration-specific parameters, which comes down to ignore any differentiated effect of macroeconomic trends on urban dynamics. This concerns the four above identified determinants, namely unitary commuting cost θ_j, relative[17] labor productivity l_j, amenity for households $U_j^{(0)}$ and the external advantages for investors $A_j^{(0)}$. We assume in addition that the intra-urban organization remains identical to its base year characteristics, as captured by constancy of parameter ξ controlling the shape of the city. Second, we do not consider any interplay between global energy dynamics and local trends, which comes down to neglect the effect of energy prices on local commuting costs and, conversely, to ignore the consequences of the spatial distribution of economic activity on global energy markets. These assumptions are adopted for the sake of simplicity, because the purpose here is not to provide realistic trends of urban dynamics but to investigate the mechanisms driving the functioning of the model.

[17] the average value of labor productivity evolves according to macroeconomic aggregates, but the productivities remain proportional to their base year value among agglomerations.

Table 5.4: *Macroeconomic aggregates in the four macro-regions for forecasts over 2001-2050*

Macroeconomic aggregates		**Region**	**2001**	**2010**	**2020**	**2030**	**2040**	**2050**
Production value	*10^{12}\$*	*USA*	18.2	22.6	26.8	31.6	36.8	43.6
		Canada	1.3	2.0	2.5	3.3	3.9	4.6
		Europe	17.1	22.9	28.0	34.9	37.6	40.6
		OECD Pacific	9.1	12.2	14.9	17.6	18.6	19.4
Effective labor	*Million workers*	*USA*	142.6	161.9	165.7	161.5	161.9	165.6
		Canada	16.0	16.8	19.5	20.3	20.7	21.1
		Europe	299.4	287.4	294.8	293.7	282.3	267.3
		OECD Pacific	105.5	100.1	104.1	100.3	93.7	85.8
Average wage rate	*10^3\$/y*	*USA*	44.5	53.1	59.1	67.2	79.0	95.3
		Canada	23.3	29.8	43.7	61.5	75.0	88.5
		Europe	14.8	17.7	23.4	32.8	37.9	42.4
		OECD Pacific	25.0	31.0	44.7	60.1	69.3	78.7
$\overline{Pop(t)}$	*millions*	*USA*	285.3	304.6	328.2	344.2	351.0	355.1
		Canada	31.1	33.0	35.2	36.6	37.2	37.5
		Europe	588.2	596.6	600.2	597.6	587.8	574.0
		OECD Pacific	204.7	208.7	208.2	204.0	198.6	190.2

We focus here only on the example of Canada, because the reduced number of agglomerations facilitates the presentation and helps identify the general mechanisms driving the dynamics of the model.

The dynamic mechanisms of urban development are driven by the attractiveness index A_j affecting firm migration decisions: agglomerations with higher attractiveness benefits from higher growth rates in market size (see figure 5.4). The number of active firms n_j in turn influences the supply side of the market, as measured by a variation in the production size of each agglomeration, n_jq_j as the amount of locally available firms varies (see figure 5.5(a)). The agglomerations with higher attractiveness then take a growing share of production.

Figure 5.4: (a) *Attractiveness index* [left-hand panel] and *(b) variations in the number of firms* [left-hand panel]

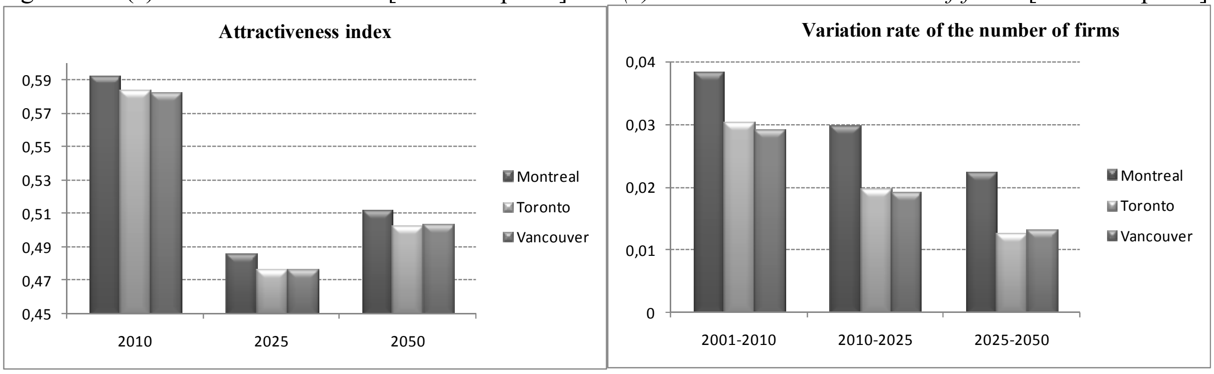
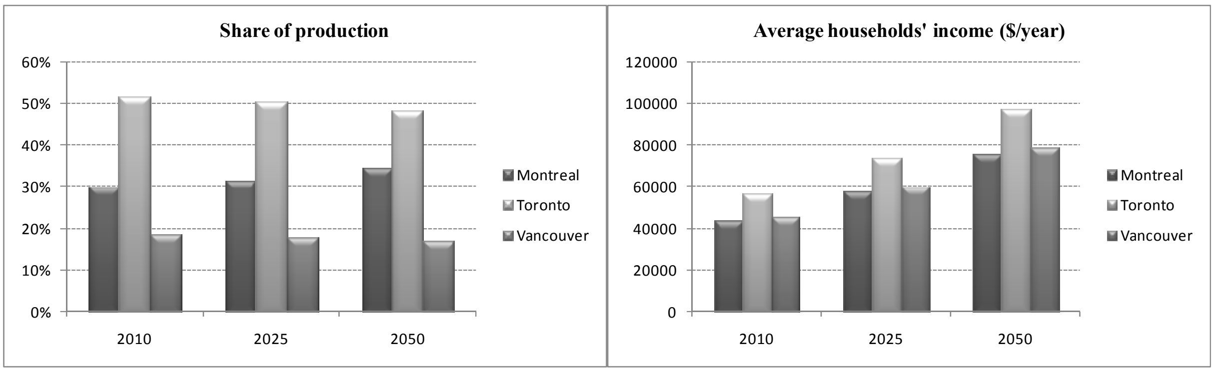


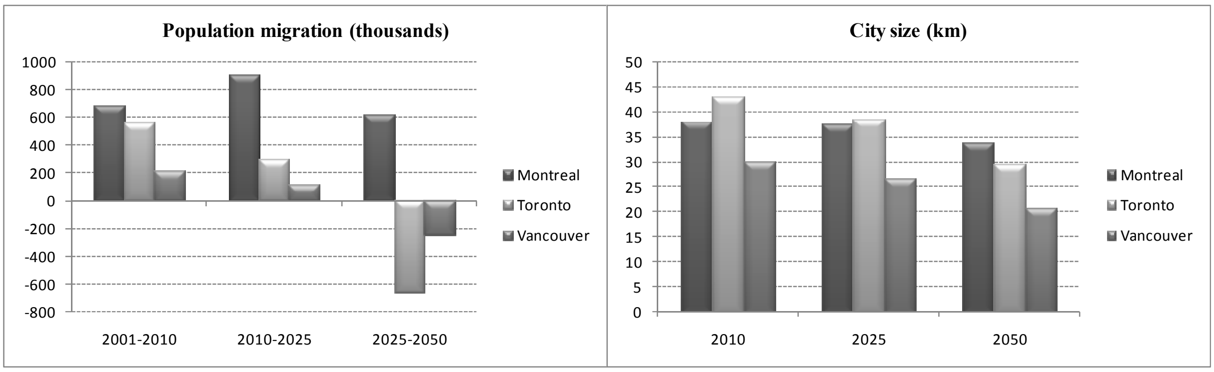
Figure 5.5: (a) *Share of production among agglomerations* [left-hand panel] and (b) *households' income* [right-hand panel]



Changes in the size of production bring about modifications of the consumption behavior of individuals located in the agglomeration j via a modification of their purchase power, as defined by $\frac{\Upsilon_j}{L_J P_j^{\beta}}$, where $\frac{\Upsilon_j}{L_J}$ represents average households' income (Figure 5.5(b)). The price index P_j, which measures the cost of living in each agglomeration, proves to remain almost homogenous across agglomerations.

Demographic trends are driven by the migration of the working force following the relocation of regional production towards the more attractive regions (Figure 5.6(a)). The consequences on the spatial structure of urban economies over the time is captured through the study of the city size *dj* , which is a proxy for the extension of the urban land area (Figure 5.6(b)).

Figure 5.6: (a) *Population migration* [left-hand panel] and *(b) city size* [right-hand panel]



VII-Conclusion

This paper develops a modeling framework in light of the New Economic Geography for long term projections of socio-economic variables at the urban scale. To improve the empirical validity of those approaches, the standard 2-region framework is extended to a multiplicity of regions in interaction, each region representing an agglomeration. The economic functioning within such region is described according to the principles of urban economy analyses, representing households' location decisions as a tradeoff between commuting and housing costs.

The model is calibrated on a set of empirical equations giving population, spatial extension, production, wage level and commuting costs for 74 of the largest OECD agglomerations. The dynamics is triggered by differences in attractiveness among regions, which capture the interplay between five determinants of agglomeration. Beyond the standard 'market size', 'cost-of-living' and 'market crowding' effects, this framework introduces the 'market density' and 'urban cost' effects. The former corresponds to the positive impact of spillover effects at the industry level permitted by an increase of the number of firms in the form of lower production costs, whereas the latter captures the constraints imposed by land availability within an agglomeration in the form of commuting and housing costs.

The analysis of long-term patterns under constant macroeconomic forcing demonstrates the possibility of partial equilibria with a distribution of economic activity among several agglomerations. The resulting 'bifurcation diagram' extends the standard core-periphery structure towards more empirically valid description of economic agglomerations. Finally, the model is submitted to changes in the macroeconomic forcing triggered by projections of economic growth drivers. This means that we investigate the effect of the macroeconomic activity on local economies, but exclude the feedback of urban dynamics on macroeconomic activity. The result is a picture of economic activities in the 74 OECD agglomerations under consideration at a 2050 horizon, with a consistent description of local activity in terms of demography, production and consumption under changing global conditions.

This analysis paves the way for quantitative long-term forecasts of consistent population and production trends at the urban level. This requires (a) embarking more detailed information on local economies and policies likely to affect agglomeration-specific development patterns and (b) considering the feedback effects of urban dynamics on global macroeconomic trends through the role of location choices on aggregate mobility needs and the investment requirements for urban infrastructures.

REFERENCES

Ciccone A (2002). Agglomeration Effects in Europe. European Economic Review 46: 213: 227.

Dimaranan, BV. (2006). *Global Trade, Assistance, and Production: The GTAP 6 Data Base*, Center for Global Trade Analysis, Purdue University.

Dixit AK and Stiglitz J. (1977). Monopolistic Competition and Optimum Product Diversity. American Economic Review 67: 297–308.

Fujita M. and Mori T (2005). Frontiers of the New Economic Geography, Papers in Regional Science 84(3): 377-407.

Krugman PR (1991). Increasing Returns and Economic Geography. Journal of Political Economy 99: 483–499.

Murata Y and Thisse JF (2005). A simple model of economic geography à la Helpman-Tabuchi. Journal of Urban Economics 58 :137-155.

Samuelson PA (1952). The Transfer Problem and Transport Costs: The Terms of Trade when the Impediments are Absent. Economic Journal 62: 278–304.

Von Thunen JH (1826). Der Isolierte Staaat in Beziehung auf Landtschaft und Nationalokonomie, Hamburg. (English translation by C.M. Wartenburg: von Thunen's Isolated State, Oxford, Pergamon Press)

United Nations (2007). World Urbanization Prospects: The 2007 Revision

United Nations (2005). World Population Prospects: The 2005 revision.

Chapter 6

Urban dynamics and climate policy: Disentangling the interplay between carbon, oil and urban land

This final chapter[*] brings together the pieces collected during this thesis to investigate the interplay between urban dynamics and the macroeconomy of low-carbon futures. This is done by organizing a dialogue between the model of urban dynamics developed in Chapter 5 and the IMACLIM-R model used in Chapter 1, 2 and 3 for the analysis of the interplay between carbon prices, oil prices and the macroeconomy. This coupled framework represents urban dynamics in changing macroeconomic conditions but also its feedback effect on long-term trends through transport and investment needs. We investigate the effect of a climate policy on urban dynamics in terms of spatial extension and housing prices (representing the value paid for the use of urban land). Conversely, we capture the influence of the urban transition on mitigation costs. In particular, additional urban infrastructure investments, while moderately reducing losses of economic activity (GDP), significantly soften welfare reductions imposed by the carbon constraint (households' surplus).

[*] A former version of this paper is given by: Grazi F, Waisman H (2009). Urban Agglomeration Economies in Climate Policy: A Dynamic CGE Approach. WP CIRED 2009-17

There is an increasing attention in climate policy literature towards the necessity of investigating the impacts of economic activities where they specifically arise (e.g., IPCC, 2001; Tietenberg, 2003; OECD, 2006; Grazi et al., 2008). Urbanization, city growth, and economic development are strictly related phenomena that have many important implications on climate change through emissions and energy consumption from economic activities located in cities. Indeed, the spatial distribution of economic activities in urban areas and its counterpart in terms of mobility needs (commuting, intra- and inter-industry trade) is an important driver of carbon emissions; conversely, a context of rising cost of fossil fuels affects the tradeoff between transport and housing prices and hence the spatial organization of cities. These interrelations suggest extending the standard focus of climate policy analysis on carbon prices to investigate the mitigation potentials associated to the interplay between spatial patterns at the urban/regional scale and carbon emissions.

However, the overwhelming majority of energy-economy models conventionnally used to assess mitigation costs focus essentially on the technological determinants of energy trends but do not capture explicitly the role of urban dynamics.[1] We propose a step forward the representation of the interplay between energy consumption, carbon emissions and the spatial organisation of cities. This is done in a framework that represents both the impact of macroeconomic trends on urban dynamics and the consequences of urban/regional location choices on mobility needs, investments and production possibilities affecting macroeconomic trajectories and carbon emissions (Section I).

We use this integrated description of global macroeconomic trajectories and location decisions in multiple agglomerations to carry out three types of analyses. First, we analyze the urban trends of a baseline scenario in which high short-term demand and resource constraints drive high long-term energy prices (section II). Second, we consider the effect of a climate policy on urban dynamics and we reassess the macroeconomic and welfare costs of a climate policy when considering the urban dimension. We focus more specifically on constrained mobility needs corresponding to daily commuting distances, investment constraints due to housing/transport infrastructure costs and housing prices representing the price paid to reside inside the city (section III). Third, we investigate the potentials offered by specific urban policies in addition to carbon pricing in an attempt to build a mix of policies apt to reduce the costs of the climate constraint. We demonstrate that the benefits of additional urban

[1] see recent exercises compiled in (IPCC, 2007).

investments are rather moderate in macroeconomic terms (GDP, carbon tax), but significant in terms of welfare (households' surplus) (section IV).

I. Modeling the interplay between urban dynamics and the macroeconomy

Numerical analysis techniques like computable general equilibrium (CGE) models developed for long-run forecasting of complex dynamic systems are seen as more reliable guides to address the relationship between determinants of economic development and forces inducing climatic variations (Böhringer and Löschel, 2006). They are based on multi-regional, multi-sectoral frameworks describing how the world economy and the adjustments of production and consumption under counterfactual scenarios representing different visions of the world or policy intervention. They give insights on the economic impacts arising from specific policy interventions by comparison of different policy measures aimed at CO2 abatement (mainly, carbon taxes/subsidies and emission trading permits) in terms of economy's efficiency, distributional effects and the cost (benefit) pressure exercised on its sectors by a tax (subsidy).

The development of those frameworks has been mostly oriented towards the representation of the technical dimension of the economy (supply-side technologies, structural change). To this aim, in line with the 'Elephant and rabbit stew metaphor' which legitimates to treat the energy sector independently from the rest of the economy because of its small share in aggregate economic activity (Hogan and Manne 1977), the efforts have been focused on the details of the representation of the energy sector. This has led in particular to the emergence of hybrid frameworks combining an explicit representation of energy technologies informed by expert-based assessments (bottom-up approach) and macroeconomic consistency in terms of money flows (top-down approach). These tools help to investigate the role of energy efficiency and technology availability when carbon price is introduced as the driver of decarbonizing economies.

But, these assessments remain at a global level and do not represent the local dimensions of interaction, so that a firm basis for thinking the interplay between carbon emissions, energy trends and the spatial organisation of economic activities in cities is still missing. We propose a step in this direction by embarking, in a hybrid energy-economy model, a representation of location choices among multiple agglomerations and within urban areas. Section I.1 describes the models used to analyse energy-economy interactions and urban systems, whereas section I.2 enters more into the details of the coupling process allowing a dialogue between them.

1- The models

1.1 The energy-economy model IMACLIM-R

We adopt the Computable General Equilibrium (CGE) model IMACLIM-R that is designed to assess the impact of energy and climate policies on carbon emissions and their cost on the economy. It contributes to the literature on energy-economy modeling by considering growth patterns in second best worlds, which are often ignored in standard models. A complete description of the IMACLIM-R model and its general characteristics in terms of climate policy analysis are given in Annex A and Chapter 3. Here, we focus on the specificities of the model that make it suitable to analyze the interplay between location patterns and macroeconomic trajectories:

- The IMACLIM-R model represents endogenous technical change determined by the volume and structure of cumulated investments given capital availability constraints. This allows in particular representing crowding-out effects on investments, and hence capturing the feedback of local infrastructure decisions at a global scale through capital availability.
- The IMACLIM-R model relies on a consistent description of the economy in money values and physical quantities in "hybrid matrixes". In addition to the preservation of energy quantities, which is commonly admitted as crucial for analysis of climate policy (Malcolm and Truong, 1999; Sands et al., 2005), this dual description is extended to cover transportation, as a prerequisite to investigate explicitly the macroeconomic effects of location choices through transport needs.[2]
- The IMACLIM-R model represents the effect of imperfect foresight: at each date, agents have limited information about the future and form their anticipations on the extrapolation of past and current trends (adaptive anticipations). This structure allows representing the inertias on agents' preferences and on the renewal of long-lived

[2] The 'hybrid matrix" at the initial date (2001) is constructed by modifying input-output tables from the GTAP-6 dataset (aggregated according to the Imaclim-R mapping in 12 regions and 12 sectors) to make them fully compatible with 2001 energy balances from IEA (in Mtoe) and passenger mobility (in passenger-km) from (Schafer and Victor, 2000).

infrastructure defining the urban structures, and the imperfect expectations about the long-term economic context due to the uncertainty on the spatial behavior when multiple locations are available.

- Finally, a specific effort has been done to represent the role of infrastructure in driving transport decisions by households. To this aim, utility-maximizing households are submitted not only to the standard income budget, but also to a time budget constraint, which limits the total time spent in transportation and drives modal allocation among four modes (automobile, public transport, air transport and non-motorized). Investments in transport infrastructure determine the efficiency of the different transport modes (i.e., the effective speed given saturation of available infrastructures) and, hence, the allocation of travel time budget across them.

1.2 The model of urban systems

The geographical dimension of economic activities is traditionally investigated by relying on analytical approaches stemming from the New Economic Geography, in line with the seminal paper by Krugman (1991). These approaches are successful in the representation of heterogeneous land uses by introducing the benefits and costs of agglomeration as driving forces of migration decisions. But, although negative environmental externalities have been included in extensions of the basic model (see discussions in Chapter 4), the stylized analytical models constitutive of the NEG approach fall short of rendering a complete picture of the complexity that animates the interplay between location choices and CO2 emissions. To overcome this limitation, we adopt a model of urban economies that extends the NEG principles to provide numerical analysis of location choices among multiple agglomerations and within urban areas, in line with both empirical evidence and microeconomic theory (Chapter 5).

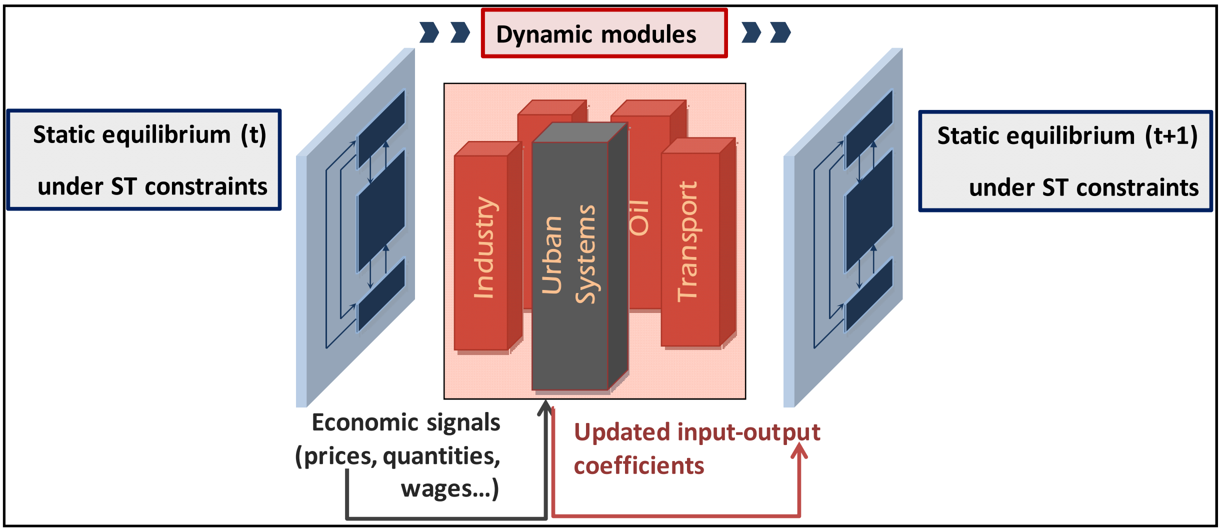
A national economy as a mass of N_A+1 regions, where N_A represents the number of agglomerations (or urban regions) under consideration and the latter is the rural area. Each agglomeration $j \in [1\ N_A]$ comprises n_j firms located in the Central Business District, which produce q_j units of a variety of a composite good with two input factors: labor (with unitary requirements l_j paid at wage rate w_j) and capital (paid at a rate of return r_j). The equilibrium price p_j is determined by monopolistic competition *à la* Dixit-Stiglitz (1977). The L_j

households are distributed within circular peripheral areas around the CBD where jobs are located, and experience commuting costs proportional to their distance x from the CBD with unitary commuting costs given by θ_j. These commuting costs affect negatively total effective labor S_j supplied by workers: $S_j < L_j$. In the rural area, land is considered as an homogenous space and the L_A households are strictly identical. Firms produce q_F units of a homogenous good under constant returns to scale with two input factors: labor (with unitary requirements l_F and wage rate w_F) and capital. The equilibrium price p_F is determined by perfect competition. The full set of equations describing the general equilibrium of this system of regional economies in interaction through trade is given by equations (1)-(49) in Chapter 5.

2- The coupled model

The IMACLIM-R model adopts recursive and modular structure, which allows incorporating sectoral expertise in dynamic modules and operationalizing the dialogue with macroeconomic trajectories through a systematic exchange of information with the static equilibrium (Figure 6.1). We exploit this possibility to represent the interplay between macroeconomic trajectories and urban dynamics by embedding the model of urban systems as a dynamic module of Imaclim-R.

Figure 6.1: *Urban systems and the modular structure of IMACLIM-R*



This means that, at each time step t, the urban module receives information from the static equilibrium in the form of aggregate macroeconomic variables, and then sends backs updated input-output coefficients resulting from changes in the urban structure to calculate the static equilibrium at date *t+1*. Due to data availability constraints, we implement this dialogue for the four OECD regions (USA, Canada, Europe and OECD Pacific) for which detailed information necessary to build the model of urban systems is fully available for the largest agglomerations.

2.1 Spatial disaggregation of macroeconomic activity

At date *t*, the model of urban systems receives information from previous static equilibrium, in the form of the average value of major macroeconomic variables at the regional level, which serve as boundary conditions for the model of urban dynamics. This first step of the dynamic interaction then consists in disaggregating the macroeconomic settings by revealing the system of urban economies that is consistent with the aggregate macroeconomic equilibrium. This means imposing that the average value of each spatially disaggregated variable equals the value of the corresponding aggregated variable at the regional level from the IMACLIM-R equilibrium.

By introducing $\overline{V(t)}$ the total value of production, $\overline{S(t)}$ the total working force (effective labor), $\overline{w(t)}$ the wage rate and $\overline{L(t)}$ the national population coming from the static equilibrium at date *t*, the equations ensuring consistency between macroeconomic and local variables are given by:

$$\sum_{k=1}^{N_A} p_k n_k q_k + p_F q_F = \overline{V(t)} \tag{1}$$

$$\sum_{k=1}^{N_A} S_k + l_F q_F = \overline{S(t)} \tag{2}$$

$$\sum_{k=1}^{N_A} w_k S_k + w_F l_F q_F = \overline{w(t)}\,\overline{S(t)} \tag{3}$$

$$\sum_{k=1}^{N_A} L_k + L_A = \overline{L(t)} \tag{4}$$

In addition to these equations controlling the size of production, labor, income and production at the national level, we consider the dependence of the two crucial agglomeration-specific determinants of urban economies to macroeconomic trajectories, namely unitary commuting costs θ_j and labor productivity l_j.

The losses of income experienced by households because of commuting amount to $\theta_j w_j$ per kilometer. We introduce a dependence of those losses to the cost of energy for transportation, in turn related to the price of liquid fuels, $\overline{p_{LF}(t)}$, and the unitary consumption of vehicles (per kilometer), $\overline{\alpha_V(t)}$, both calculated in previous static equilibrium. More precisely, we consider that commuting costs at the agglomeration level evolve in time proportionally to the unitary cost of transportation at the national level, $\overline{p_{LF}(t)}.\overline{\alpha_V(t)}$. This can be written as:

$$\theta_j(t) = \theta_j(t-1).\left[\frac{\overline{p_{LF}(t)}}{\overline{p_{LF}(t-1)}} + \frac{\overline{\alpha_V(t)}}{\overline{\alpha_V(t-1)}} - \frac{w_j(t)}{w_j(t-1)} - 1\right] + 1 \tag{5}$$

Macroeconomic trajectories consider aggregate labor productivity growth following a convergence hypothesis (Barro and Sala-i-Martin, 1992), the parameters being calibrated on historic trajectories (Maddison, 1995) and 'educated guess' assumptions of long-term trends (Oliveira-Martins et al., 2005): USA remains the world leader in productivity per worker with a steady growth of 1.65% per year, whereas the dynamics of productivity in other countries is driven by a partial catch-up (the lower the absolute productivity per worker in a country, the higher its labor productivity growth). For the sake of simplicity, we assume that these gains of productivity are uniformly distributed among the agglomerations, so that the relative gains of local productivity are identical in all regional economies. By noting $\overline{l(t)}$ the unitary requirement for production at the national level (ie. the inverse of labor productivity), we have:

$$l_j(t) = l_j(t-1).\frac{\overline{l(t)}}{\overline{l(t-1)}} \tag{6}$$

For each region of the IMACLIM-R model, the first step of the dynamic model formally comes down solving the system of equations defining urban economies (equations (1)-(49) in Chapter 5) under the constraint of equations (1)-(6) defining the boundary conditions imposed by macroeconomic trends from the static equilibrium of the IMACLIM-R model. This ensures a spatial disaggregation of the economy into a set of urban agglomerations (and a rural area).

2.2 Urban dynamics and the feedback effect on macroeconomic trends

The dynamics of the model of urban economies is enabled through firms' location decisions taken in function of differences in attractiveness for investments among agglomerations, as measured by the return to capital r_j. The resulting changes in the spatial distribution of production and productive capital in the national economy bring about migrations of households/workers to satisfy constraints on each local labor market. At the agglomeration level, the spatial distribution of households inside each agglomeration is decided according to a minimization of infrastructure and congestion costs.

These dynamic mechanisms impose modifications of some important economic variables at the agglomeration level, which, when re-aggregated at the national level, result in modified constraints on the subsequent static equilibrium. In this exercise, we focus on the feedback effect of urban dynamics on four crucial determinants of the macroeconomic equilibrium: basic needs of transport, transport capacities, investments in urban infrastructure and labor productivity.

(a) The basic needs of transport represent constrained mobility, ie the volume of transport that households have no choice but to do. We identify this constrained mobility with commuting distance, as given by the distance to the CBD in urban agglomerations. By introducing the land consumption by each household, $\lambda_j(x)$, the total of commuting distances D_j by all households residing in a given agglomeration j is given by

$$D_j = \int_{-d_j}^{d_j} \frac{x}{\lambda_j(x)} dx \tag{7}$$

Under the assumption of a power functional form $\lambda_j(x) = \overline{\lambda_j} x^{\xi}$, this can be rewritten as

$$D_j = 2 \frac{d_j^{\,2-\xi}}{\overline{\lambda_j}\left(2-\xi\right)} \tag{8}$$

The basic needs for transport at the national level, *bn*, are then given by the sum of commuting distances over all agglomerations:

$$bn = \sum_{j=1}^{N_A} D_j \tag{9}$$

(b) The density of urban agglomerations also affects the modal distribution of transport capacities, which measures the availability of transport infrastructures. We consider indeed that denser agglomerations favor the deployment of public transportation at the expense of automobile. To capture this effect, we define the average density of urban settlements ρ as the ratio between urban population and the spatial extension of all urban areas:

$$\rho = \frac{\sum_{j=1}^{N_A} L_j}{\sum_{j=1}^{N_A} d_j} \tag{10}$$

The amount of transport capacities in public transportation, Cap_{public}, is then an increasing function of ρ, whereas transport capacities for private vehicles (urban roads), Cap_{cars}, is a decreasing function of ρ. For the sake of simplicity, we assume these dependences to be proportional to changes in density.

(c) The needs for investments in urban infrastructure depend on the density of settlements inside agglomerations in a way to capture higher marginal construction costs in the building sector and the need for more developed transport infrastructure in denser cities. By introducing $\gamma > 1$ as a measure of the non-linearity of urban capital requirements, the amount of investment per capita I increases with density as given by the inverse of the land consumption by each household, $\lambda_j(x)$:

$$I_j(x) = I_0 \left(\frac{1}{\lambda_j(x)} \right)^{\gamma - 1} \tag{11}$$

where I_0 normalizes the units of measurement. The total amount of investments mobilized in infrastructure in the j-agglomeration, IC_j, is then given by:

$$IC_j = 2\int_0^{d_j} I_0 \left(\frac{1}{\lambda_j x^{\xi}} \right)^{\gamma} dx = \frac{2.I_0}{\lambda_j^{\gamma}} \frac{d_j^{1-\gamma\xi}}{1-\gamma\xi} \tag{12}$$

The total of investments at the national level, *DI*, is in turn defined by the sum of investments over all agglomerations

$$DI = \sum_{j=1}^{N_A} IC_j \tag{13}$$

These investments are financed by the government, and ultimately affect households' income by reducing public transfers.

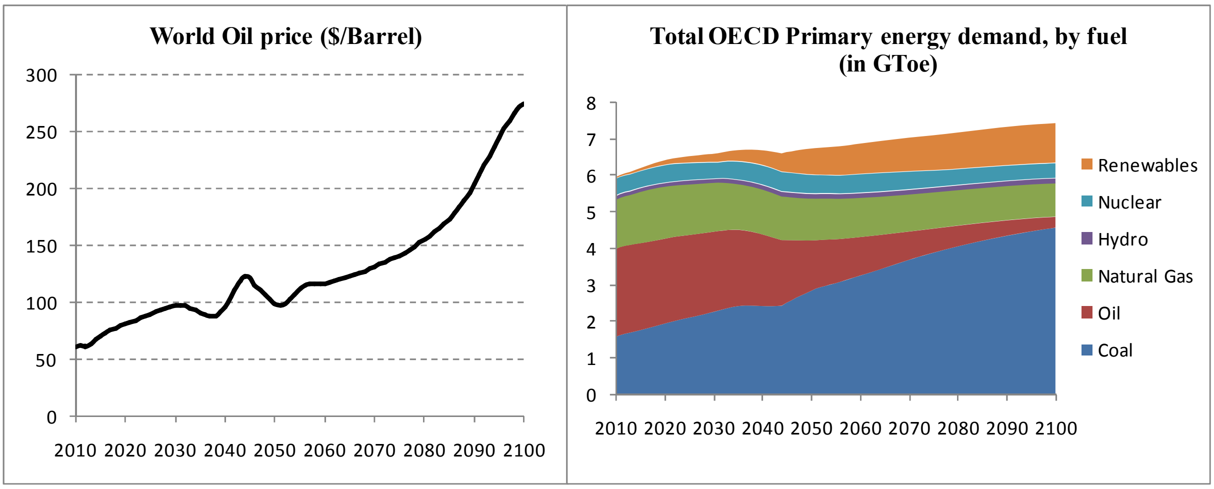
(d) Finally, the relocation of production among agglomerations of different productivities l_j implies changes in the average productivity at the national level. The relative change in productivity $\frac{\Delta l}{l}$ resulting from firms' migration decisions is given by:

$$\frac{\Delta l}{l} = \frac{\dfrac{\sum_{j=1}^{N_A} n_j(t+1)^{1-\alpha} l_j q_j}{\sum_{j=1}^{N_A} n_j(t+1) q_j}}{\dfrac{\sum_{j=1}^{N_A} n_j(t)^{1-\alpha} l_j q_j}{\sum_{j=1}^{N_A} n_j(t) q_j}} \tag{14}$$

II. Consistent urban trends behind the macrodynamics of the baseline scenario

In absence of climate policy, demographic trends in OECD countries feature a small decrease of total population over 2010-2100 (from 1.15 to 1.07 Billion persons), but productivity gains are sufficient to ensure a steady increase of economic activity at an average growth rate of 1.34% over 2010-2100. Total primary energy demand increases from 5.94 GTep in 2010 to 7.45GTep in 2100, this rather moderate increase being permitted by annual energy efficiency gains of 1.08%. This period is marked by the depletion of oil reserves, which results in a sharp increase of oil prices (Figure 6.2(a)) and a progressive switch in the energy mix towards coal, which is the more abundant fossil energy, and renewable energies (Figure 6.2(b)). In this scenario, OECD carbon emissions increase from 15.5GtCO_2 in 2010 to 21.5GtCO_2 in 2100 in parallel with the diffusion of coal liquefaction as the major substitute to oil for liquid fuel production in the second half of the century.

Figure 6.2: *(a) World oil price ($/Barrel)* and *(b) Primary energy demand (Gtoe)*



The coupled model allows giving the urbanization trends that are consistent with the above described macroeconomic trends in each of the four OECD regions (USA, Canada, Europe, OECD Pacific). In a given region, the distribution of production between the major urban agglomerations and the "rural area" (which encompasses small and medium-size cities and dispersed settlements) is driven by firms' migration decisions and triggered by differences of attractiveness for productive investments. These endogenous mechanisms ultimately determine the redistribution of population between the "rural area" and the largest agglomerations, as summarized by the share of total population that resides in the major agglomerations (Table 6.1).

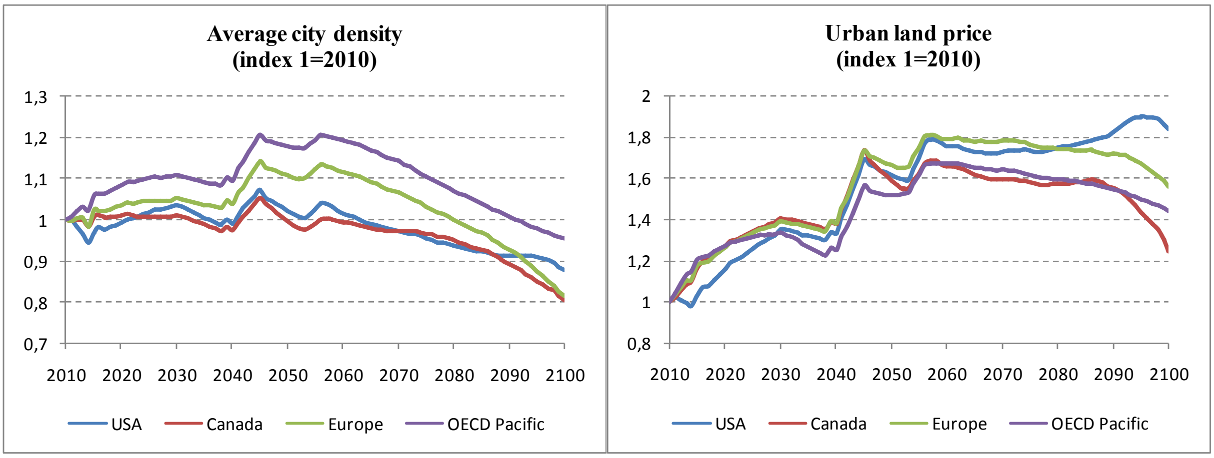
Table 6.1: *Share of total population in the largest urban agglomerations (%)*

	2010	*2020*	*2030*	*2050*	*2100*
USA	39.1	39.4	39.0	37.3	34.4
CANADA	36.5	35.6	34.8	32.2	31.7
EUROPE	25.2	24.7	24.4	22.9	22.3
OECD PACIFIC	51.2	49.5	48.5	45.4	41.1

To interpret correctly this indicator, two points are worth noting. First, its absolute value depends on the range of agglomerations adopted in this study, which correspond to those with more than 1 Million habitants at the base year. The importance of those agglomerations in total population strongly varies across regions, as captured by differences in the initial share of "urban" population, from 25% in Europe to 51% in OECD Pacific. Second, the decreasing trend of the share does not necessary reflect massive migrations towards rural zones, but rather a redispersion of population towards smaller urban units.

Changes in the distribution of production among agglomerations also drive the evolution of population living in each location, since households constitute the working force that is necessary to ensure local production. This relation between production and population is even more direct under the assumption that all households must live in the very agglomeration where they have their work. To characterize the spatial distribution of population within agglomerations, we consider the average urban density and average urban land price (Figure 6.3).[3]

Figure 6.3: *(a) City density* and *(b) Urban land price*



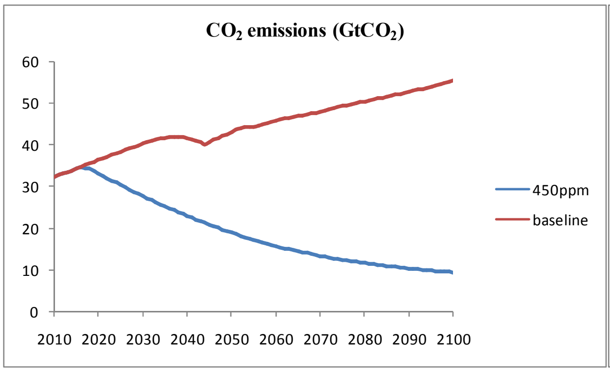
The general urban trends are logically similar in all regions, as a reaction to energy price variations and efficiency gains. The first half of the century features a stabilization of density at a slightly higher level than at its base year level as a consequence of energy price increases affecting transport costs and hence fostering denser settlements and an increase of urban land prices. After 2050, the diffusion of coal-to-liquid as an abundant substitute to oil helps maintaining a moderate price of liquid fuels despite the rise of oil prices and the cumulative effect of energy efficiency gains is sufficient to ensure a decrease of unitary commuting costs. This fosters urban sprawl and a moderation of urban land prices in more dispersed settlements.

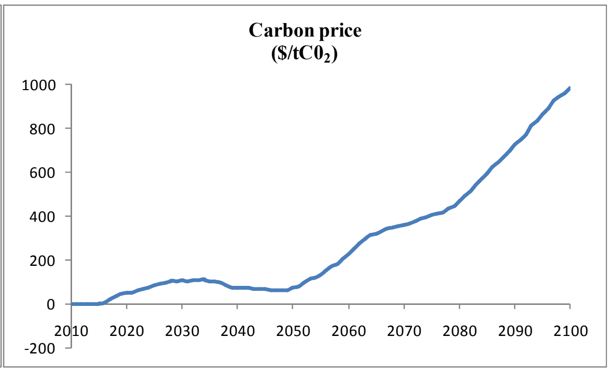
[3]The average value is calculated as a weighted mean of the values over all agglomerations, in each OECD region.

III- Climate policy and urban systems

We now turn to the implementation of an ambitious climate policy modeled, for the sake of simplicity, by a prescribed path for carbon emissions (Figure 6.4(a)). The targeted emission trajectory is chosen in category II of IPCC scenarios corresponding to a stabilization target of 440-485 ppm CO_2: global CO_2 emissions peak in 2017 and are decreased by 20% and 60% with respect of 2000 level in 2050 and 2100, respectively (IPCC, 2007, Table SPM5). The model endogenously calculates the carbon tax to be imposed at each point in time to satisfy this emissions trajectory (Figure 6.4(b)). For the sake of simplicity, we assume that the carbon tax is uniform across sectors, households and regions and exclude international redistribution of tax revenues.

Figure 6.4: *(a) CO_2 emissions in the baseline and climate policy scenarios (G tCO_2); (b) Carbon price (\$/tCO_2).*

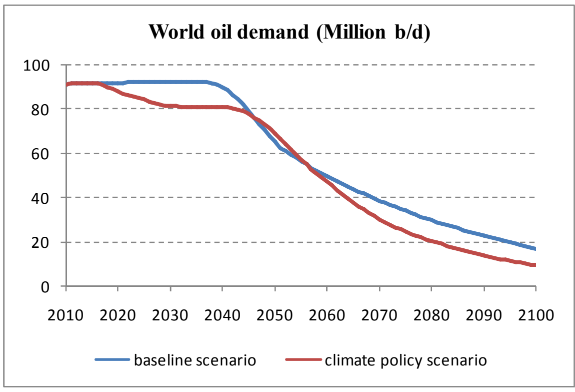


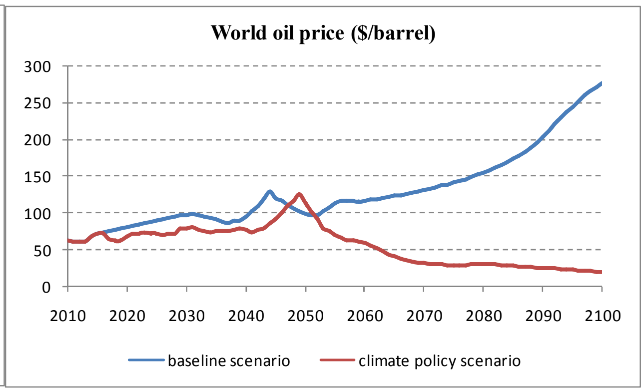


Three phases of carbon price dynamics can be identified (see Chapter 3 for a more detailed discussion). In the short-term, the carbon price increases rather sharply and reaches 100\$/tCO_2 around 2030, because of the necessity to give a strong early signal to trigger emission reductions despite inertias on the renewal of technical systems and imperfect foresight. In the medium-term (2030-2060), the carbon price tends to decline, since carbon prices above 50\$/tCO_2 are sufficient to reach most of mitigation potentials in industry, residential and power sectors, which form the core of emission reductions. In the long-term, the carbon price features a sharp increase in order to reach the remaining high-cost mitigation potentials (especially in the transport sector).[4]

[4] Note that this effect is even more pronounced, since we have adopted conservative assumptions on the diffusion of oil substitutes, like biofuels and Electric Vehicles. The sensitivity of the costs of the climate policy on these assumptions and a more in-depth analysis of the three phases of mitigation costs are given in Chapter 3

Figure 6.5: *(a) World oil demand (MBarrel/day); (b) World oil price ($/barrel)*





The climate policy affects oil markets through an acceleration of oil-free technical change in the short-term and a reduction of the long-term dependence on oil (Figure 6.5(a)). This means in particular that the climate policy makes oil-importing economies less vulnerable to oil disruptions as captured by continuously lower oil prices than in the baseline case until 2050 (Figure 6.5(b)).[5] The long-term difference is particularly pronounced with complete divergence of price trajectories during the final phase of oil depletion (after 2060, total production is lower than 50 MBarrel/day). The high long-term oil prices in the baseline case are the symptom of an economy that remains rather dependent on oil and is then affected by the decrease of production capacities. On the contrary, under climate policy, cumulated low-fossil technical change permits a lower dependence on oil, as measured by a 45% lower demand in 2100 than in the baseline case, so that supply capacities remain sufficient and depletion constraints are not binding.

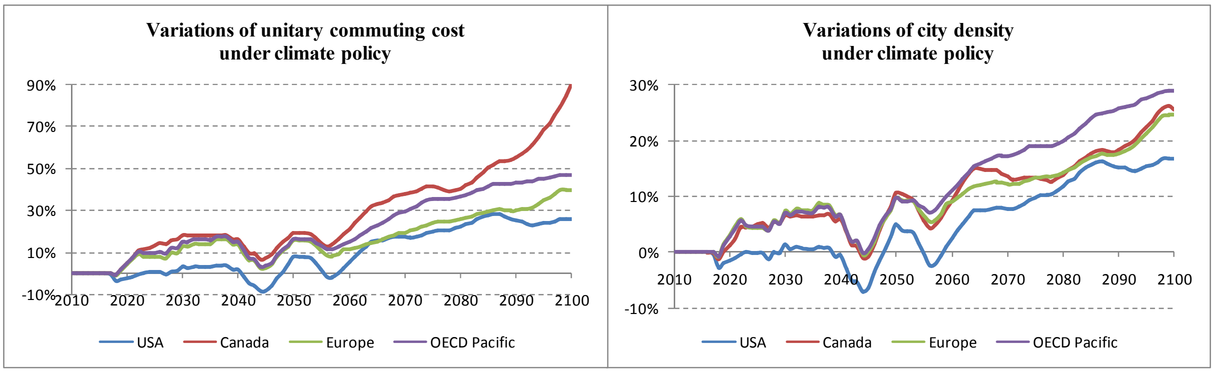
These changes affect urban dynamics through their effect on the cost of mobility in urban areas, as measured by the variations of unitary commuting costs (per km) due to the climate policy (Figure 6.6(a)). Until 2050, this variable remains rather close to its baseline levels, reflecting a moderate increase of the price of liquid fuels, since the additional carbon cost is partially compensated by the lowering of oil prices. During the second half of the century, unitary commuting costs are notably increased by the climate policy because of the concomitance of (i) a stagnation of long-term productivity and efficiency gains on vehicles and (ii) a sharp rise of carbon prices triggering high liquid fuel cost[6]. The increase of unitary commuting costs acts as an incentive for limiting the dependence on transport and hence

[5] (Rozenberg et al, 2010) investigates more in-depth this effect, and Chapter 2 analyzes the consequences of these adverse effects of the climate policy from oil exporters' point of view.

[6] liquid fuels production remains dependent on fossil energy, since one important substitute to oil is coal liquefaction, an even more-carbon intensive source of energy

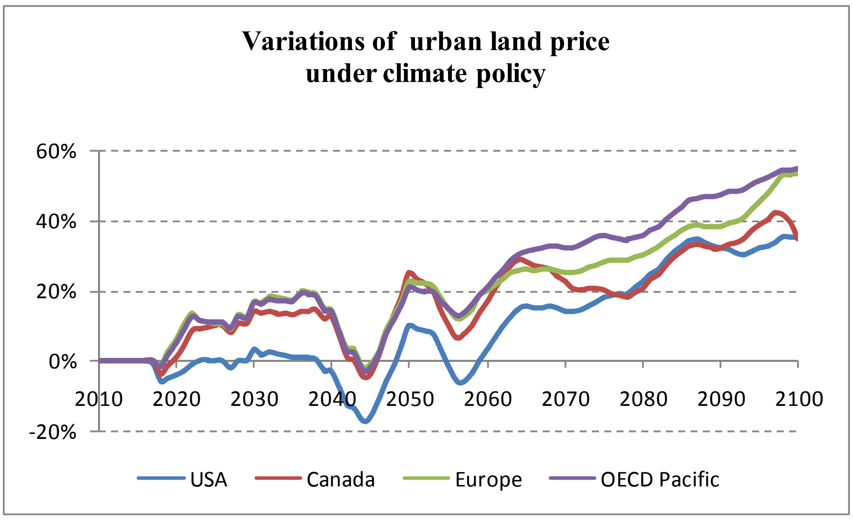
adopting denser urban centers, as captured by the increase of average urban density (Figure 6.6(b)).

Figure 6.6: *Variations under climate policy of (a) the unitary commuting cost; (b) the urban density.*



This densification process supposes a redirection of investments towards the infrastructure sector, which receives an annual average of 12 additional Billion$ to support the cost of capital-intensive investments in dense centers (high buildings, dense public transport). Since competition for land is more intense in denser settlements, the densification process fostered by the climate policy goes along with a rise of land prices (Figure 6.7). This effect is particularly important in the second half of the century, with a 30-50% rise of land prices in 2100.

Figure 6.7: *Variations of urban land price under climate policy*

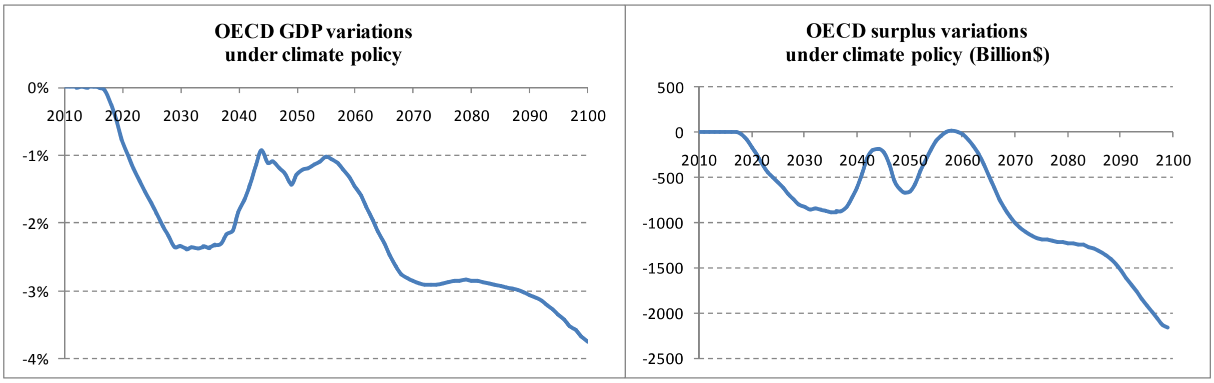


The global economic effects of the climate policy on the production side are measured by GDP changes (Figure 6.8(a)), whereas surplus variations[7] capture the welfare effects on

[7] The surplus variations measure the amount of money that should be given to leave households' utility identical in the two scenarios (see discussion in Chapters 1 and 2). Note that the surplus variations in Figure 6.8(b) are

households (Figure 6.8(b)). In line with the above analysis of the three phases of carbon prices, we obtain significant short-term economic and welfare losses, a medium-term partial GDP recovery during which households catch-up their baseline utility level and a final drop of economic and welfare indicators.

Figure 6.8: *Variations under climate policy of (a) OECD GDP (%); and (b) OECD Surplus (Billion$).*



IV- The contribution of urban planning to climate policy

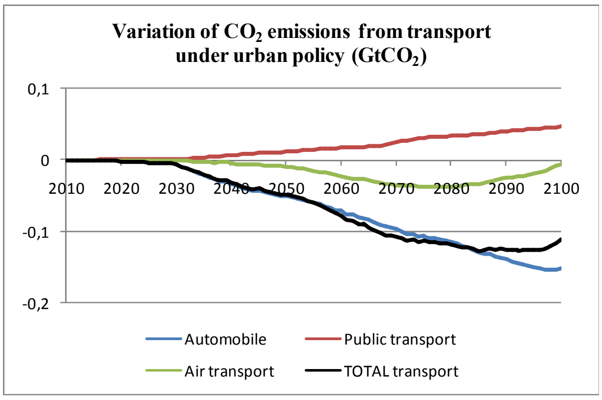
The above analysis demonstrates that the economic and welfare costs due to a climate policy remain important if it relies essentially on the implementation of a carbon price, especially after 2060. This section considers a broader architecture, in which this global approach is complemented by more local measures aimed at anticipating the increase of commuting costs fostering urban densification under climate policy. More specifically, we consider policies that consist in mobilizing investments for urban infrastructures in order to accelerate the changes towards more appropriate denser urban structures in a climate policy context. We consider that, after progressively entering in force, the total amount of annual investments (deducted from governments' budgets) reaches 0.1% of OECD GDP and serves to accelerate the densification of urban areas.

This urban policy decreases the dependence on mobility, and then contributes to decrease carbon emission from the transport sector by 0.1 $GtCO_2$ in 2100 (Figure 6.9). The reductions of carbon emissions essentially come from the automobile sector, which is the dominant commuting mode and hence the more affected by the changes of urban structures. More indirectly, the denser agglomerations offer more importance to public transport and hence a

calculated under the assumption that housing goods are incorporated in a composite aggregate and hence treated like any other good. This approach neglects some specificity of housing goods in households' utility due to their limited substituability with consumption goods and the strong inertia constraints on the supply-side, notably due to urban infrastructure dynamics.

rebound of its emissions, while the promotion of low-speed modes affects indirectly air transport activity through the time budget constraint.

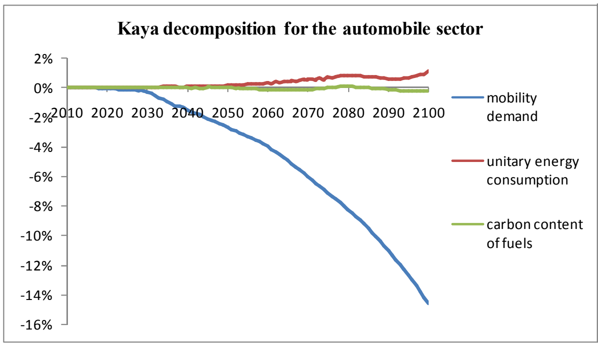
Figure 6.9: *Variations under urban policy of transport CO_2 emissions*



To investigate more in-depth the variations of carbon emissions from the automobile sector, we perform a 'Kaya' decomposition in order to identify the relative roles of mobility demand (in p-km), vehicles' energy intensity (in l/km) and the carbon content of fuels (in gCO_2/km):

$$Emissions = \underbrace{\frac{Emissions}{Energy}}_{carbon\ content} \cdot \underbrace{\frac{Energy}{pkm}}_{energy\ intensity} \cdot \underbrace{pkm}_{volume\ effect}$$

Figure 6.10:*(a) Kaya decomposition of the automobile sector;(b) variations of world oil price under urban policy*



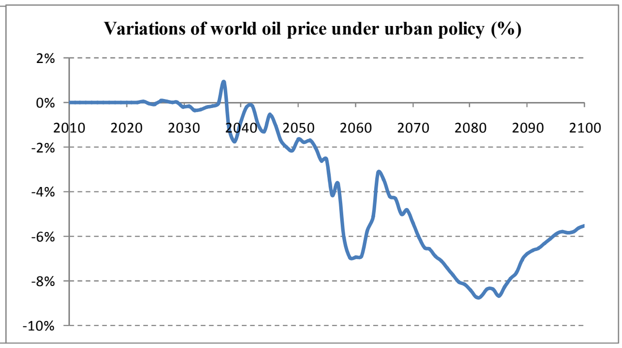
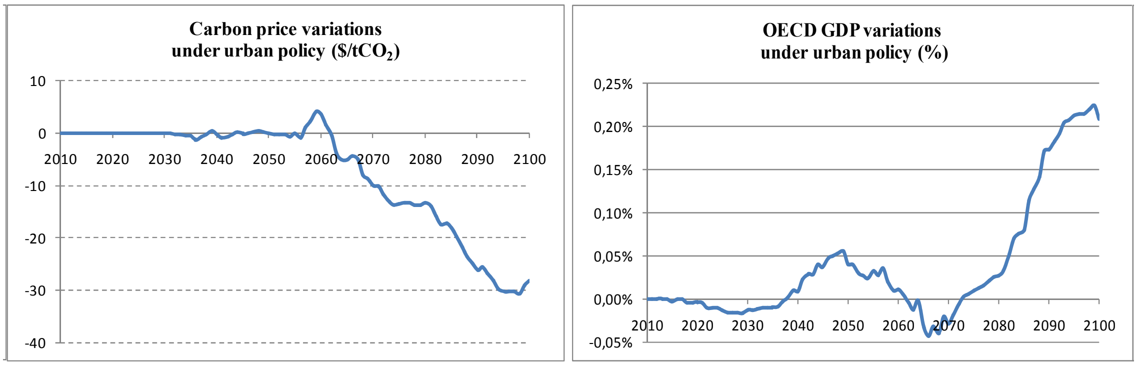


Figure 6.10(a) shows relative variations of the three components of the Kaya identity when the urban policy is set in place. Expectedly, the major impact of the urban policy occurs through a reduction of the total volume of transport (measured in passenger-km) due to a decrease in average commuting distance by individuals in denser agglomerations (up to 14% reduction in 2100). However, indirect effects also simultaneously occur that affect carbon intensity and vehicle unitary fuel consumption in the long run. In particular, the decrease of liquid fuel demand due to lower commuting distances endogenously generates a fall of liquid

fuel prices (Figure 6.10(b)). This drop in turn slows down technical change towards more energy-efficient vehicles, and automobile vehicles are 1.1% less efficient in 2100 under urban policy.

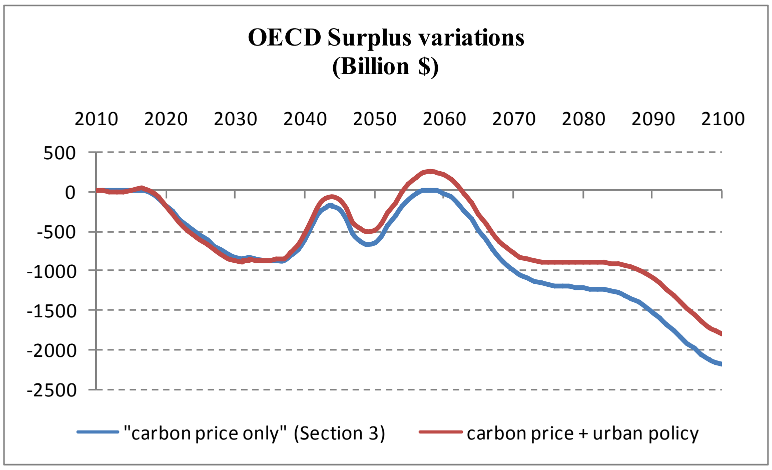
The 0.1 GtCO2 represent 8% of total emission reductions in the transport sector at this time horizon under climate policy and the urban policy then facilitates the climate policy, as captured by a relative decrease of the carbon tax to be imposed for the same stabilization objective (Figure 6.11a). However, this direct effect of the urban policy on carbon emissions remains moderate with regard to the global effort. This is confirmed by the analysis of global economic activity, which features a moderate increase of GDP with respect to the "carbon price only" policy (up to 0.2% in 2100), especially when compared to the total mitigation costs (-3.7% in 2100) (Figure 6.11b).

Figure 6.11: *Variations under urban policy of (a) carbon price ($/tCO2); (b) OECD GDP (%)*



This evaluation changes when considering surplus variations (Figure 6.12): the implementation of urban policy helps reducing OECD surplus losses by 18% in 2100, with a notable effect from 2050.

Figure 6.12: *OECD Surplus variations under climate policy (Billion $)*



The magnitude of this effect can be understood because of the interplay between urban dynamics and constrained mobility associated to daily commuting, which acts as a crucial constraint for households and undermines their purchase power. In a context of climate policy fostering high fuel costs, this constrained mobility is particularly expensive and explains the important welfare losses in the long-term. Under additional infrastructure investments towards, denser urban settlements reduce this constrained mobility and promote low-cost public transport modes, hence easing the burden imposed by mobility needs on households' income and ultimately improving their purchase power.

V- Carbon price, oil price, land price: the drivers of macroeconomic costs

Here, we synthesize the results obtained in section III and IV, which correspond to different architectures of climate policies, where carbon pricing is either the only policy adopted for carbon emissions reductions or is complemented by local urban infrastructure policies. Our analysis demonstrates that these two architectures correspond to different sets of carbon/oil/land prices with different effects on the cost of climate policies. To represent the OECD tradeoffs at different time horizons, we synthesize the results for two discount rate values, 7% representing short-sighted decisions and 1% long-term focused approaches (Table 6.2). [8]

Unsurprisingly, urban policies have hardly any effect in the short-term (discount rate 7%), because of inertias on the deployment of infrastructures. Considering a 1% discount rate, the set of prices is modified by the implementation of the urban policy, which fosters a 2% decrease of carbon and oil prices versus a 13% increase of urban land prices. This latter trend is associated with denser settlements, which foster a reduction of mobility constraints. These changes have limited effect on GDP losses but affect importantly surplus losses, which are reduced by 20% for the same stabilization objective.

[8] These values of the discount rates are adopted to provide an appraisal of economic effects at different time horizons. A 7% value boils down to considering essentially the effects during the first 20 years (the weighting factor is lower than 0.25 for all years after 2030); on the contrary, with a 1% discount rate, one "sees" the trajectory over the whole period 2010-2100 since the weighting factor is always higher than 0.4 even in 2100.

Table 6.2: Economic and welfare indicators under different climate policy architectures

	discount rate = 7%		discount rate = 1%	
	carbon price only	carbon price & urban policy	carbon price only	carbon price & urban policy
Carbon price *($/tCO$_2$)*	56.2	55.8	225.0	219.8
Oil price *($/Barrel)*	69.4	69.2	61.2	60.0
Land price *(index 1 =baseline)*	1.31	1.37	1.70	1.93
GDP variations *(Trillion $)*	- 6.24	- 6.23	- 68.9	- 67.5
Aggregate surplus variations *(Trillion $)*	- 4.27	- 4.20	- 39.1	- 31.1

VI- Conclusion

This paper has presented an innovative modeling framework for a consistent analysis of the feedback mechanisms between urban, regional (country) and global (world) economies in the context of climate change. This is done by embarking a model of urban systems disaggregating production, consumption and trade among multiple cities into the recursive dynamic, computable general-equilibrium model IMACLIM-R. This offers an innovative framework, which differs from earlier work by allowing a spatially disaggregated investigation instead of the global approach conventionally adopted for climate policy analysis.

This model is used to analyze the effects of macroeconomic trends, oil price trajectories and technical change on the long-term dynamics of urban systems. We demonstrate in particular that the rise of energy prices increases commuting costs in urban areas and hence fosters a urban densification process during the next decades, whereas the diffusion of energy efficiency decreases the commuting costs and favors a dispersion of urban settlements in the long term. When a climate policy is adopted, the higher cost of fossil fuels triggers a densification process in urban areas, which tend to increase the welfare losses due to the climate policy.

We finally test the implementation of explicit policies at the urban scale, in the form of investments devoted to improve the compatibility of urban systems to the climate policy. These urban policies prove to reduce the cost of climate policy by facilitating the decrease of transport-related carbon emissions, but have the indirect effect of forcing further densification with losses of housing welfare. The overall effect however remains positive and such local policy adopted at the urban scale ensures a 20% decrease of welfare losses. Here, it is worth recalling that the magnitude of the effects can be undestood as a lower bound since (i) the urban policies envisaged involve only OECD agglomerations at the exclusion of major urban cities in the developing world and (ii) the amount of investments devoted to these policies is limited to 0.1% of OECD GDP

REFERENCES

Barro RJ and Sala-i-Martin X (1992). Convergence Journal of Political Economy 100(2): 223-251

Böhringer C and Löschel A (2006). Computable General Equilibrium Models for Sustainability Impact Assessment: Status Quo and Prospects. Ecological Economics 60: 49–64.

Dixit AK and Stiglitz JE (1977). Monopolistic Competition and Optimum Product Diversity. Amer. Econ. Rev. 67: 297–308.

Grazi, F., J.C.J.M. van den Bergh, and J.N. van Ommeren. 2008. An Empirical Analysis of Urban Form, Transport, and Global Warming. The Energy Journal 29(4): 97-107.

Hogan, W.W. and A.S. Manne (1977). "Energy-Economy Interactions: The Fable of the Elephant and the Rabbit?," in C.J. Hitch (Ed), *Modeling Energy-Economy Interactions: Five Approaches*: 247-277. Resources for the Future, Washington D.C.

IPCC (2007). Summary for Policymakers. In: B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (eds) Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. CambridgeUniversity Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2001). Climate Change 2001: The Scientifc Basis. D.L. Albritton et al. (eds.). Cambridge University Press, New York.

Krugman PR (1991). Increasing Returns and Economic Geography. J. Polit. Economy 99: 483–499.

Maddison A (1995) Monitoring the World Economy: 1820 – 1992 *OECD Development Center, 260 p*

Malcolm G and Truong P (1999). The Process of Incorporating Energy Data into GTAP, *Draft GTAP Technical Paper*, Center for Global Trade Analysis, Purdue University, WestLafayette, Indiana, USA.

OECD (2006).Competitive cities in the global economy.. OECD Territorial Revies. Organisation for Economic Co-operation and Development, Paris.

Oliveira Martins J, Gonand F, Antolin P, de la Maisonneuve C and Kwang-Y (2005). The impact of ageing on demand, factor markets and growth *OECD Economics Department Working Papers*, #420, OECD Economics Department

Rozenberg J, Hallegatte S, Vogt-Schilb A, Sassi O, Guivarch C, Waisman H and Hourcade JC. (2010). Climate policies as a hedge against the uncertainty on future oil supply, *Climatic Change (forthcoming)*

Sands RD, Miller S and Kim MK (2005). The Second Generation Model: Comparison of SGM and GTAP Approaches to Data Development, *Pacific Northwest National Labouratory*, PNNL-15467, 2005.

Schäfer A and Victor DG, (2000). The future mobility of future population .*Transportation Research Part A*, 34: 171-205

Tietenberg TH (2003) .The tradable permits approach to protecting the commons: lessons for climate change. Oxford Review of Economic Policy 19: 400-419.

General Conclusion

This thesis contributes to the on-going effort of assessment of the consequences of climate policies in support to political negotiations towards a climate agreement. It is more particularly focused on the constraints imposed by scarce natural resources, which are mostly ignored by existing analyses but play a crucial role in the context of climate policies. This is particularly true for oil and urban land use, which are both crucial drivers of emissions from the transport sector. As noted in the introduction, the effect of accounting for the limitations imposed by these natural resources is *a priori* ambiguous since it suggests both the risk of enhanced costs if carbon limitations reinforce the sub-optimalities caused by pre-existing constraints, but also, conversely, the possibility of co-benefits if the climate policy helps to correct some pre-existing imperfections.

The answers to these questions are built upon the elaboration of a modeling framework able to capture the specificities of oil and urban land uses and their interplay with global macroeconomic interactions. This has been done through the coupling of the general equilibrium framework IMACLIM, designed to represent energy-economy interactions for climate measures assessment, with two innovative models capturing the determinants of oil supply constraints and location decisions among a system of cities in interaction:

- The model of oil supply constraints accounts for geological, geopolitical and economic constraints on the deployment of oil production capacities. This is done by representing explictly different categories of oil, distinguished by their cost of exploration/exploitation, and by imposing, for each of them, a yearly maximum rate of increase capturing depletion constraints. Non-Middle East producers deploy all production capacities that are profitable given selling price at their maximum rate, whereas Middle-East producers can adapt freely their production capacities within these limits in function of their price objective.
- The model of urban agglomerations in interaction represents location decisions among a system of several agglomerations and within urban areas in function of economic tradeoffs. The agglomerations are in interaction through interregional trade and firms' migration decisions which can decide to move from one location to another according to profitability prospects. Households' location decisions within urban areas result from a tradeoff between commuting and housing costs, in astylized vision of a monocentric, axysymmetric city.

Our findings can be summarized in three points:

- The date of Peak Oil is sensitive to short-term oil price only in case of high reserves but this date is not necessarily informative for the time profiles of oil prices, rent formation and growth patterns. In particular, inertia and imperfect foresight create the possibility of a sudden and potentially long lasting acceleration of oil price increases at the Peak Oil Period. This effect is amplified if the economy is very oil-dependent like after long periods of low energy prices and creates the possibility of an important economic crisis. On the contrary, low oil price in the short-term may be in the interests of oil producers bacause it fosters a steady increase of exportation revenues over the long term in absence of climate policy
- The macroeconomic losses due to a climate policy for major oil exporters are rather important in the short term but moderate in the long term. Despite significant reductions of long-term exportation revenues, structural readjustment towards local industrialisation instead of oil exportations proves benefitial in the long-term in a formalization of the "Dutch Disease" mechanisms. At the global level, the cost of climate policies is driven by the ratio energy-to-labor costs, which measures the energy intensity of the economy. Technical inertias under imperfect foresight and reduced vulnerability to Peak Oil (energy security effect) explain the important short-term losses and the medium-term partial recovery. At a long term horizon, the major driver is the mitigation potentials in the transport sector, either through low-carbon options or reduced mobility needs through spatial reorganisation.
- Location decisions depend on a tradeoff between centripetal forces fostering agglomeration (market size, diversity of goods, economies of scale) and centrifugal forces fostering dispersion of activities (congestion, environmental pollution). Empirically, urban dynamic trends are well reproduced when assuming that firms' relocation decisions are driven by the return to capital. Urban infrastructure policies aimed at controlling urban mobility through spatial reorganisation have almost no effect as a complementary measure to carbon pricing in the short term because of inertias on the deployment of infrastructures. In the long-term, on the contrary, they help to reduce significantly the welfare costs of climate policies thanks to lower carbon and energy prices but this comes at the cost of an increase of urban land prices.

These analyses suggest some policy messages on important dimensions of climate negotiations:

- Oil-importing economies are vulnerable to Peak Oil and face the risk of a deep crisis if the oil dependency is not reduced significantly. It may thus be in their interest to correct potentially misleading price-signals by using complementary measures to secure steady technical change. Among them, climate policies help to soften the cost of the oil transition by accelerating fossil-free technical change and can act as a hedging strategy against the uncertainty on oil supply.
- The monetary compensations to oil exporters for the adverse impacts of climate policies are unacceptable if based on losses of oil exportation revenues, but they become more debatable if losses of economic activity are considered. Indeed, in this latter case, the transfers pass through a peak but then decline over time to be almost zero in 2050.
- In the short-term, the curbing of carbon emissions make necessary a steep increase of the carbon price signal to accelerate technical change depsite inertias and imperfect expectations. This causes important transitory losses, especially in developing regions, if implemented with a carbon tax and lump-sum recycling. This calls for the adoption of complementary measures on labor market to soften this transition through, e.g., carbon revenue recycling
- Over the long run, the major issue concerns the transport sector and important losses may be experienced in the pessimistic case of low availability of carbon-free vehicles (like Electric Vehicles) and continuation of current trends in terms of mobility. Because of the weak sensitivity of the transport activity to energy price signals, this highlights the importance of complementary policies to carbon pricing designed to tackle transport-related emissions through a combination of incentives for technical progress on vehicles and changes in the spatial organization helping to reduce mobility needs.
- Urban policies adopted in climate policy frameworks to limit the increase of constrained mobility need have indirect effect on investment availability (crowding-out effects), energy prices and urban land prices, which are essential to provide a complete assessment of the desirability of such measures. The analysis proves that, when taking into accoutnall these effects, a moderate mobilization of investments for

specific urban policies (around 0.1% of GDP) helps to significantly reduce the long-term welfare losses

The modeling framework developed in this thesis constitutes a methodological advancement that allows widening the discussions about climate policy to the investigation of issues raised by urban dimensions under carbon constraints. This work opens the way for further research in echo with modeling developments aimed at improving the representation of some specific aspects of the three spatial scales in interaction:

- At the local scale, the priority will be to improve the representation of urban infrastructure to reach a more precise description of location decisions and transport choices than with the current stylized description of a monocentric city. This means in particular representing the possibility of polycentric settlements and explicit modal choice for constrained mobility in line with available transport infrastructures.
- At the regional scale, the major limitation of inter-city interactions concerns the description of trade barriers, which are currently modeled with a 'iceberg' structure instead of an explicit transport sector. To overcome this shortcoming, the model will be extended to represent exlicitly freight transport associated with the spatail organization of economic activity and populations, as well as inter-agglomeration passenger mobility.
- At the global scale, the major advancement to be undertaken concerns the representation of macroeoconomic effects associated with the circulation of land rents in terms of investments. This means in particular relaxing the conventional assumption of equilibrated balance of payments to consider alternative assumptions on governments' budget and international capital flows. This means also revisiting the way investments are allocated among sectors to represent the possibility of positive effects permitted by rent captation, which could help redirecting investment towards more efficient uses compared to the options adopted by rent-seekers.

Annexes

Annex A

This technical appendix complements the descriptions of the IMACLIM-R model used in Chapters 1, 2, 3 and 6 by providing detailed information on modeling assumptions in the static equilibrium (section I), on the Nexus describing technical change in the energy sector (section II), on data defining the calibration date and 'natural' growth drivers (Section III), the regional and sectoral disaggregation (Section IV). In Section V, we discuss the assumptions and calculations supporting the analytical analysis of the drivers of mitigation costs used in Chapter 3.

I- Variables, parameters and equations of the static equilibrium

We distinguish between endogenous variables (marked in bold) and fixed parameters of the static equilibrium at date *t*. For the sake of readability, indexes *i* and *j* are used for sectors, and index *k* is reserved for regions

1. Table of variables

Table A-1 details the list of variables calculated by the static equilibrium.

Table A-1: *Variables of the static equilibrium.*

$\mathbf{Income_k}$	Households' total revenues in region k
$\mathbf{transfers_k}$	Transfers from States to households in region k
$\mathbf{p_{k,i}}$	Production price of good i in region k
$\mathbf{pC_{k,i}}$	Final consumption price for households for good i in region k
$\mathbf{pG_{k,i}}$	Final consumption price for States for good i in region k
$\mathbf{pI_{k,i}}$	Price for investments for good i in region k
$\mathbf{pIC_{j,i,k}}$	Intermediate consumption price for sector i for good j in region k
$\mathbf{pind_k}$	Households final consumption price index in region k
$\mathbf{wp_i}$	International price of good i
$p_{k,i}^{imp}$	Import price of good i in region k
$\mathbf{w_{k,i}}$	Unitary salary in sector i in region k
$\mathbf{\Omega_{k,i}}$	Increasing cost factor in sector i in region k
$\mathbf{Q_{k,i}}$	Volume of production of good i in region k
$\mathbf{C_{k,i}}$	Households final consumption volume of good i in region k
$\boldsymbol{S_{k,mobility}}$	Households' demand for mobility services
$\mathbf{pkm_{k,mode}}$	Passengers.kilometers travelled per mode (air transport, public transport, private vehicle, non motorized mode) in region k
$\mathbf{I_{k,i}}$	Volume of good i purchased for Gross Fixed Capital Formation (Investment) in region k
$\mathbf{z_k}$	Unemployment level in region k
$\mathbf{M_{k,i}}$	Volume of imports of good i in region k
$\mathbf{X_{k,i}}$	Volume of exports of good i from region k
$\mathbf{X_i}$	Volume of the international market of good i

Table A-1(continued): *Variables of the static equilibrium.*

$MS^{X}_{k,i}$	Market share of exports from region k in the international market of good i
$shareC^{imp/dom}_{k,i}$	Imports (/Domestic production) share in households final consumption of good i in region k
$shareG^{imp/dom}_{k,i}$	Imports (/Domestic production) share in States final consumption of good i in region k
$shareI^{imp/dom}_{k,i}$	Imports (/Domestic production) share in investments of good i in region k
$shareIC^{imp/dom}_{k,i}$	Imports (/Domestic production) share in sector i intermediate consumption of good j in region k
$\mathbf{NRB_k}$	Net regional savings of region k
$\mathbf{GRB_k}$	Gross regional savings of region k
$\mathbf{InvFin_{k,i}}$	Investment allocated to sector i in region k
$\mathbf{pCap_{k,i}}$	Price of one unit of productive capital in sector i and region k
$\mathbf{\Delta Cap_{k,i}}$	New productive capital in sector i and region k

2- Table of parameters

Table A-2 details the parameters, which are fixed in each static equilibrium and are modified in the recursive framework by dynamic modules

Table A-2: *Parameters of the static equilibria.*

$G_{k,i}$	States final consumption of good i in region k
$IC_{j,i,k}$	Sector i intermediate consumption of good i in region k
L_k	Total active population in region k
$l_{k,i}$	Quantity of labour per unit of output in sector i in region k
$aw_{k,i}$	Wage curve parameter for sector i in region k
$\pi_{k,i}$	Markup rate in sector i in region k
ptc_k	Households propensity to spend (one minus saving rate) in region k
$div_{k,i}$	Share of profits in sector i in region k given as revenues to households

Table A-2(continued): *Parameters of the static equilibria.*

$bn_{k,i}$	Basic need of consumption of good i in region k
$\alpha_{k,Ei}^{cars}$	Mean consumption of energy Ei per passenger.kilometer by car in region k
$\alpha_{k,Ei}^{m2}$	Mean consumption of energy Ei per square meter of residential buildings in region k
$Tdisp_{k}$	Total households travel time in region k
$Cap_{k,i}$	Productive capacity of sector i in region k
$Captransport_{k,j}$	Total capacity of transport mode j in region k
$tax_{k,i}^{w}$	Labour tax rate in sector i in region k
$tax_{k,i}^{M}$	Tax rate on imports of good i in region k
$tax_{k,i}^{X}$	Tax rate on exports of good i from region k
$tax_{k,i}^{domC}$	Tax rate on households final consumption of domestic production of good i in region k
$tax_{k,i}^{impC}$	Tax rate on households final consumption of imports of good i in region k
$shareExpK_{k}$	Share of gross regional savings of region k exported to the international 'pool' of capital
$shareImpK_{k}$	Share of the international 'pool' of capital imported in region k
$shareInvFin_{k,i}$	Share of net regional savings of region k allocated to sector i
$\beta_{j,i,k}$	Quantity of good j necessary to build one unit of productive capacity of sector i in region k
$nit_{k,i}^{it}$	Transport need in mode it for imports of good i in region k
$\xi_{k,i}^{C}$, $\xi_{k,i}^{S}$	Parameters of the utility function
$b_{k,mode}$	Calibration parameters for the constant elasticity of substitution function giving the transport service in function of passengers.kilometers per mode in region k
η_{k}	$\eta = \frac{s-1}{s}$, with s the elasticity of substitution of the function giving the transport service in function of passengers.kilometers per mode in region k
$wref_{k,i}$	Salaries at calibration date in sector i in region k
$pindref_{k}$	Households final consumption price index in region k at calibration date

Table A-2(continued): *Parameters of the static equilibria.*

$zref_k$	Underutilization of the labour force at the calibration date for region k
$\rho_{k,i}$	$\rho = \frac{1-\sigma}{\sigma}$
$\sigma_{k,i}$	Armington elasticity for good i in region k
b^{dom}, b^{imp}	Calibration parameters for Armington expression for good i in region k
θ_i	$\theta = \frac{1-\lambda}{\lambda}$
λ_i	Armington elasticity in the international market for good i
$\psi_{k,i}$	Calibration parameter for Armington expression for exports of good i from region k in the international market 'pool'
$\eta_{k,i}^{imp}$	Parameter for the expression of the imports (/Domestic production) share in households final consumption of good i in region k
$\eta_{k,i}^{X}$	Parameter for the expression of the market share of exports from region k in the international market of good i

3. Core equations of the static equilibrium

Income formation

$$\mathbf{Income_k} = \sum_{sectors\ i} \mathbf{\Omega_{k,i}} \cdot \mathbf{w_{k,i}} \cdot l_{k,i} \cdot \mathbf{Q_{k,i}} + \sum_{sectors\ i} div_{k,i} \cdot \pi_{k,i} \cdot \mathbf{p_{k,i}} \cdot \mathbf{Q_{k,i}} + \mathbf{transfers_k} \quad (1)$$

Governments' budget

$$\sum \mathbf{taxes} = \sum_{sectors\ i} G_{k,i} \cdot \mathbf{pG_{k,i}} + \mathbf{transfers_k} + InvInfra_k$$

The sum of taxes corresponds to the total of tax revenues, i.e. the tax rates (parameters) applied to the taxable amounts (often endogenous in the equilibrium).

Utility maximisation

$$\mathbf{U_k}\left(\vec{\mathbf{C}}_\mathbf{k}, \vec{\mathbf{S}}_\mathbf{k}\right) = \left[\prod_{\text{goods } i} \left(\mathbf{C_{k,i}} - bn_{k,i}\right)^{\xi_{k,i}^{C}}\right]\left(\mathbf{S_{k,mobility}} - bn_{k,mobility}\right)^{\xi_{k,j}^{S}} \quad (2)$$

$$\mathbf{S_{k,mobility}} = \left(\left(\frac{\mathbf{pkm_{k,air}}}{b_{k,air}} \right)^{\eta_k} + \left(\frac{\mathbf{pkm_{k,public}}}{b_{k,public}} \right)^{\eta_k} + \left(\frac{\mathbf{pkm_{k,cars}}}{b_{k,cars}} \right)^{\eta_k} + \left(\frac{\mathbf{pkm_{k,nonmotorized}}}{b_{k,nonmotorized}} \right)^{\eta_k} \right)^{\frac{1}{\eta_k}} \quad (3)$$

Income constraint

$$ptc_k \cdot \mathbf{Income}_k = \sum_{sectors\ i} \mathbf{pC_{k,i}} \cdot \mathbf{C_{k,i}} + \sum_{Energies\ Ei} \mathbf{pC_{k,Ei}} \cdot \left(\mathbf{pkm_k^{cars}} \cdot \alpha_{k,Ei}^{cars} + S_k^{m^2} \cdot \alpha_{k,Ei}^{m^2} \right) \quad (4)$$

Travel time budget constraint

$$Tdisp_k = \sum_{\text{means of transport j}} \int_0^{\mathbf{pkm_{k,j}}} \tau_{k,j} \left(\frac{u}{Captransport_{k,j}} \right) du \ , \quad (5)$$

where τ_j represents the marginal efficiency in transport time (the time necessary to travel an additional passenger.kilometer with mode j) :

$$\tau_{k,j}(x) = a_{k,j} \cdot x^{ktrans_{k,j}} + b_{k,j} \ . \quad (6)$$

The first order conditions give N+S equations, with N the number of consumption goods and S the number of mobility services, and add two unknowns, the Lagrange multipliers for both constraints.

Sector budget (supply curve)

$$\mathbf{p_{k,i}} = \sum_{sectors\ j} \mathbf{pIC_{j,i,k}} \cdot IC_{j,i,k} + \left(\mathbf{\Omega_{k,i}} \cdot \mathbf{w_{k,i}} \right) \cdot l_{k,i} \cdot (1 + tax_{k,i}^{w}) + \pi_{k,i} \cdot \mathbf{p_{k,i}} \quad (7)$$

$\mathbf{\Omega_{k,i}} = \Omega \left(\frac{\mathbf{Q_{k,i}}}{Cap_{k,i}} \right)$ represents an increasing cost (or decreasing returns) function of the productive capacities utilisation rate. The functional form for Ω is:

$$\mathbf{\Omega_{k,i}} = a_\Omega - b_\Omega \cdot \tanh\left(c_\Omega \cdot \left(1 - \frac{Q}{Cap} \right) \right) \quad (8)$$

Labor market (wage curve)

$$\mathbf{z_k} = 1 - \frac{\sum_{sectors\ i} l_{k,i} \cdot \mathbf{Q_{k,i}}}{L_k} \tag{9}$$

$$\frac{\boldsymbol{w_{k,i}}}{\boldsymbol{pind_k}} = aw_{k,i} \cdot \frac{wref_{k,i}}{pindref_k} \cdot f\left(\frac{\mathbf{z_k}}{zref_k}\right) \tag{10}$$

Equilibrium constraints on physical flows

$$\mathbf{M_{k,i}} = \mathbf{shareC_{k,i}^{imp}} \cdot \mathbf{C_{k,i}} + \mathbf{shareG_{k,i}^{imp}} \cdot G_{k,i} + \mathbf{shareI_{k,i}^{imp}} \cdot \mathbf{I_{k,i}} + [\sum_{sectors\ j} \mathbf{Q_{k,j}} \cdot IC_{i,j,k}^{imp} \cdot \mathbf{shareIC_{i,j,k}^{imp}}] \tag{11}$$

$$\mathbf{Q_{k,i}} = \mathbf{shareC_{k,i}^{dom}} \cdot \mathbf{C_{k,i}} + \mathbf{shareG_{k,i}^{dom}} \cdot G_{k,i} + \mathbf{shareI_{k,i}^{dom}} \cdot \mathbf{I_{k,i}} + [\sum_{sectors\ j} \mathbf{Q_{k,j}} \cdot IC_{i,j,k} \cdot \mathbf{shareIC_{i,j,k}^{dom}}] + \mathbf{X_{k,i}} \tag{12}$$

Investment formation

$$\mathbf{NRB_k} = \mathbf{GRB_k} \cdot (1 - shareExpK_k) + \left(\sum_{countries\ k'} \mathbf{GRB_{k'}} \cdot shareExpK_{k'}\right) \cdot shareImpK_k \tag{13}$$

$$\mathbf{GRB_k} = \mathbf{Income_k} \cdot \left(1 - ptc_k\right) + \sum_{sectors\ j} \pi_{k,j} \cdot \mathbf{p_{k,j}} \cdot \mathbf{Q_{k,j}} \cdot \left(1 - div_{k,j}\right) \tag{14}$$

$$\mathbf{InvFin_{k,i}} = \mathbf{NRB_k} \cdot shareInvFin_{k,i} \tag{15}$$

$$\mathbf{pCap_{k,i}} = \sum_{sectors\ j} \left(\beta_{j,i,k} \cdot \mathbf{pI_{j,i,k}} \right) \tag{16}$$

$$\mathbf{\Delta Cap_{k,i}} = \frac{\mathbf{InvFin_{k,i}}}{\mathbf{pCap_{k,i}}} \tag{17}$$

$$\mathbf{I_{k,j}} = \sum_{\text{sectors } i} \beta_{j,i,k} \cdot \mathbf{\Delta Cap_{k,i}} \tag{18}$$

4. Intermediate variables for international trade

<u>*Armington goods*</u>

$$\mathbf{C_{k,i}} = \left(b_{k,i}^{dom} \cdot \left(\mathbf{C_{k,i}^{dom}} \right)^{-\rho_{k,i}} + b_{k,i}^{imp} \cdot \left(\mathbf{C_{k,i}^{imp}} \right)^{-\rho_{k,i}} \right)^{-\frac{1}{\rho_{k,i}}} \tag{19}$$

$$\mathbf{pC_{k,i}} = \left(\left(b_{k,i}^{dom} \right)^{\sigma_{k,i}} \left(\mathbf{p_{k,i}} \cdot (1 + tax_{k,i}^{domC}) \right)^{1-\sigma_{k,i}} + \left(1 - b_{k,i}^{dom} \right)^{\sigma_{k,i}} \left(\mathbf{p_{k,i}^{imp}} \cdot (1 + tax_{k,i}^{impC}) \right)^{1-\sigma_{k,i}} \right)^{\frac{1}{1-\sigma_{k,i}}} \tag{20}$$

$$\mathbf{shareC_{k,i}^{dom}} = \left(b_{k,i}^{dom} \cdot \frac{\mathbf{pC_{k,i}}}{\mathbf{p_{k,i}} \cdot (1 + tax_{k,i}^{domC})} \right)^{\sigma_{k,i}} \tag{21}$$

$$\mathbf{shareC_{k,i}^{imp}} = \left(\left(1 - b_{k,i}^{dom} \right) \cdot \frac{\mathbf{pC_{k,i}}}{\mathbf{p_{k,i}^{imp}} \cdot (1 + tax_{k,i}^{impC})} \right)^{\sigma_{k,i}} \tag{22}$$

Similar equations to (-19)–(-22) are valid for public consumptions, investments and intermediate consumptions.

$$\mathbf{p_{k,i}^{imp}} = \mathbf{wp_i} \cdot \left(1 + taxe_{k,i}^{M} \right) + \sum_{means\ of\ tranport\ it} \mathbf{wp_{it}} \cdot nit_{k,i}^{it} \tag{23}$$

$$\sum_{countries\ k} \left(\mathbf{shareC_{k,i}^{imp}} \cdot \mathbf{C_{k,i}} + \mathbf{shareG_{k,i}^{imp}} \cdot G_{k,i} + \mathbf{shareI_{k,i}^{imp}} \cdot \mathbf{I_{k,i}} + \sum_{sectors\ j} \mathbf{shareIC_{i,j,k}^{imp}} \cdot IC_{i,j,k} \cdot \mathbf{Q_{k,j}} \right) = \mathbf{X_i} = \left[\sum_{countries\ k} \psi_{k,i} \cdot \mathbf{X_{k,i}^{-\theta_i}} \right]^{-\frac{1}{\theta_i}} \tag{24}$$

$$\mathbf{X_{k,i}} = \left[\psi_{k,i} \cdot \frac{\mathbf{wp_i}}{\mathbf{p_{k,i}} \cdot (1 + tax^{X}_{k,i})} \right]^{\lambda_i} \cdot \mathbf{X_i} \tag{25}$$

$$\mathbf{wp_i} = \left(\sum_{countries\ k} \left(\psi_{k,i}\right)^{\lambda_i} \left(\mathbf{p_{k,i}} \cdot (1 + tax^{X}_{k,i}) \right)^{1-\lambda_i} \right)^{\frac{1}{1-\lambda_i}} \tag{26}$$

<u>Energy goods</u>

$$\mathbf{C_{k,i}} = \mathbf{C^{dom}_{k,i}} + \mathbf{C^{imp}_{k,i}} \tag{27}$$

$$\mathbf{pC_{k,i}} = \mathbf{shareC^{dom}_{k,i}} \cdot \mathbf{p_{k,i}} \cdot (1 + tax^{domC}_{k,i}) + \mathbf{shareC^{imp}_{k,i}} \cdot \mathbf{p^{imp}_{k,i}} \cdot (1 + tax^{impC}_{k,i}) \tag{28}$$

$$\mathbf{shareC^{imp}_{k,i}}(\mathbf{t}) = \frac{shareC^{imp}_{k,i}(t-1) \cdot \left(\dfrac{\mathbf{p^{imp}_{k,i}}(t) \cdot (1 + tax^{impC}_{k,i}(t))}{p^{imp}_{k,i}(t\text{-}1) \cdot (1 + tax^{impC}_{k,i}(t\text{-}1))} \right)^{\eta^{imp}_{k,i}}}{shareC^{imp}_{k,i}(t-1) \cdot \left(\dfrac{\mathbf{p^{imp}_{k,i}}(t) \cdot (1 + tax^{impC}_{k,i}(t))}{p^{imp}_{k,i}(t\text{-}1) \cdot (1 + tax^{impC}_{k,i}(t\text{-}1))} \right)^{\eta^{imp}_{k,i}} + \left(1 - shareC^{imp}_{k,i}(t\text{-}1)\right) \cdot \left(\dfrac{\mathbf{p_{k,i}}(t) \cdot (1 + tax^{domC}_{k,i}(t))}{p_{k,i}(t-1) \cdot (1 + tax^{domC}_{k,i}(t\text{-}1))} \right)^{\eta^{imp}_{k,i}}} \tag{29}$$

$$\mathbf{shareC^{dom}_{k,i}}(\mathbf{t}) = 1 - \mathbf{shareC^{imp}_{k,i}}(\mathbf{t}) \tag{30}$$

Similar equations to (27)–(30) are valid for public consumptions, investments and intermediate consumptions.

$$\mathbf{p^{imp}_{k,i}} = \mathbf{wp_i} \cdot \left(1 + taxe^{M}_{k,i}\right) + \sum_{means\ of\ tranport\ it} \mathbf{wp_{it}} \cdot nit^{it}_{k,i} \tag{31}$$

$$\sum_{countries\ k}\left(\mathbf{shareC^{imp}_{k,i}}\cdot\mathbf{C_{k,i}}+\mathbf{shareG^{imp}_{k,i}}\cdot G_{k,i}+\mathbf{shareI^{imp}_{k,i}}\cdot\mathbf{I_{k,i}}+\sum_{sectors\ j}\mathbf{shareIC^{imp}_{i,j,k}}\cdot IC_{i,j,k}\cdot\mathbf{Q_{k,j}}\right)=\mathbf{X_i}$$

(32)

$$\mathbf{MS^X_{k,i}}(\mathbf{t})=\frac{MS^X_{k,i}(t-1)\cdot\left(\frac{\mathbf{p_{k,i}}(t)\cdot\left(1+tax^X_{k,i}(t)\right)}{p_{k,i}(t-1)\cdot\left(1+tax^X_{k,i}(t-1)\right)}\right)^{\eta^X_{k,i}}}{\sum_{countries\ k'}MS^X_{k',i}(t-1)\cdot\left(\frac{\mathbf{p_{k',i}}(t)\cdot\left(1+tax^X_{k',i}(t)\right)}{p_{k',i}(t-1)\cdot\left(1+tax^X_{k',i}(t-1)\right)}\right)^{\eta^X_{k',i}}} \tag{33}$$

$$\mathbf{X_{k,i}}=\mathbf{MS^X_{k,i}}(t)\cdot\mathbf{X_i} \tag{34}$$

$$\mathbf{wp_i}=\frac{\sum_{countries\ k}\mathbf{p_{k,i}}\cdot\left(1+tax^X_{k,i}\right)\cdot\mathbf{X_{k,i}}}{\sum_{countries\ k}\mathbf{X_{k,i}}} \tag{35}$$

II- The dynamic modules of IMACLIM-R

The purpose of this section is to describe the NEXUS of Imaclim-R, which determine technical change through the evolution of production costs and end-use equipments. We start by describing the evolution of the constraints on fossil fuel production (oil, coal, gas) before turning to energy transformation (liquid fuels and electricity). Finally, we turn our attention to the technical coefficients driving final energy consumption in both stationary uses (industry and residential uses) and non-stationary uses (freight and passenger transportation).

1 Modelling primary supply of fossil fuels

1.1 Oil supply

The ‘oil supply’ Nexus embarks three crucial specificities of oil supply:
(a) a small group of suppliers benefits from a market power.

(b) the geological nature of oil reserves imposes a limited adaptability of oil supply. (c) uncertainties on the technical, geopolitical and economical determinants of oil markets alter agents' expectations. The assumption of perfectly optimizing atomistic agents, which remains a useful analytical benchmark, fails to provide a good proxy for the oil economy.

We distinguish seven categories of conventional and five categories of non-conventional oil resources in each region. Each category i is characterized by the amount of ultimate resources[1] $Q_{\infty,i}$ and by a threshold selling price above which producers initiate production, $p^{(0)}(i)$. This price is a proxy for production costs and accessibility.

Each oil category is submitted to geological constraints (inertias in the exploration process and depletion effects), which limit the pace of expansion of their production capacity. In line with (Rehrl and Friedrich, 2006), who combine analyzes of discovery processes (Uhler, 1976) and of the "mineral economy" (Reynolds, 1999), we impose, at each date *t*, an upper bound $\Delta Cap_{\max}(t,i)$ on the increase of production capacity for an oil category *i :*

$$\frac{\Delta Cap_{\max}(t,i)}{Cap(t,i)} = \frac{b_i \cdot \left(e^{-b_i(t-t_{0,i})} - 1\right)}{\left(1 + e^{-b_i(t-t_{0,i})}\right)} \tag{36}$$

The parameter b_i (in t^{-1}) controls the intensity of constraints on production growth: a all (high) b_i means a flat (sloping) production profile to represent slow (fast) deployment of production capacities. The parameter $t_{0,i}$ represents the date at which production capacities of the concerned oil category are expected to start their decline due to depletion effects. It is endogenous and varies in time since it depends on the amount of oil remaining in the soil given past exploitation decisions.

The production decisions of non-Middle-East producers are those of 'fatal producers' who do not act strategically on oil markets and invest in new production capacity if an oil category becomes profitable given the selling oil price p_{oil}. They develop production capacities at their maximum rate of increase in eq (1) for least-cost categories ($p_{oil} > p^{(0)}(i)$) but stop investments in high-cost categories ($p_{oil} < p^{(0)}(i)$). If prices continuously increase, production capacities of a given oil category follow a bell-shape trend, whereas their deployment profile passes through a plateau if prices decrease below the profitability threshold.

[1] Ultimate resource of a given category is the sum of resources extracted before 2001 and recoverable resources.

Middle-East producers are 'swing producers' who fill the gap between fatal producers' supply and global oil demand. The stagnation and decline of conventional oil in the rest of the world temporarily reinforces their market power and they can control the time profile of oil prices through the utilization rate of production capacities (Kaufmann et al, 2004). They can decide to slow the development of production capacities down (below the maximum increase given by eq (1)) in order to adjust the oil price according to their rent-seeking objective.

Total oil production capacity at date *t* is given by the sum over oil categories with different production costs (captured by different $p^{(0)}(i)$ threshold). This means that projects of various merit orders coexist at a given point in time, consistently with the observed evidence[2] and theoretical justifications[3].

1.2 Gas supply

The evolution of worldwide natural gas production capacities meets demand increase until available reserves enter a depletion process. Distribution of regional production capacities in the 'gas supply' Nexus is made using an exogenous distribution key calibrated on the output of the POLES energy model (LEPII-EPE, 2006), which captures reserve availability and regional production facilities. Gas markets follow oil markets with a 0.68 elasticity of gas to oil price. This behavior is calibrated on the World Energy Model (IEA, 2007) and is valid as long as oil prices remain below a threshold $p_{oil/gas}$. At high price levels reflecting tensions due to depletion of reserves, gas prices are driven by production costs and the increased margin for the possessors of the remaining reserves.

1.3 Coal markets

Unlike oil and gas markets, cumulated coal production has a weak influence on coal prices because of large world resources. Coal prices then depend on current production through elasticity coefficients. To represent the asymmetry in coal price response to production variations, we consider two different values of this elasticity, η^{+}_{coal} and η^{-}_{coal} , the former (latter) corresponding to a price reaction to a production increase (decrease). Tight coal

[2]for example, low-cost fields in Saudi Arabia and high-cost non-conventional production in Canada are simultaneously active on oil markets

[3](Kemp and Van Long, 1980) have indeed demonstrated that, in a general equilibrium context, the lowest-cost deposits are not necessarily exploited first. (Holland, 2003) even demonstrates that least-cost-first extraction rule does not hold in partial equilibrium under capacity constraints, like those envisaged for geological reasons here.

markets exhibit a high value of η^{+}_{coal} (i.e the coal price strongly increases if production rises) and low value of η^{-}_{coal} (the price decreases only slightly if production drops).

2. Energy transformation

2.1 Liquid fuels

The ‘Substitutes to oil ‘nexus considers two large-scale substitutes to oil for liquid fuel production.

The first large-scale substitute to oil for liquid fuels production consist in first and second generation biofuels from renewable land resources. Their diffusion is controlled by supply curves: at each date, biofuels’ market share is an increasing function of oil price, carbon tax included, $S_{bio}(t,p_{oil})$.[4] This captures, although in a simplistic manner, the competition between biofuels and oil-based liquid fuels: everything else being equal, the former are more competitive and their penetration into the market is more prominent when higher oil price make the latter more expensive. The supply curves include asymptote representing explicit limits on production due to constraints on land availability and competition with other biomass uses. They are modified from one date to the other to account for learning-by-doing improvements. The diffusion of biofuels is in addition submitted to the constraint of a time delay, Δt_{bio}, which captures inertia on the deployment of raw products (biomass) and of refining capacity.

The second alternative to oil is Coal-To-Liquid (CTL). We consider it as an inexhaustible[5] backstop technology but submitted to capacity constraints. In line with Amigues et al (2001), production of the inexhaustible substitute starts before all the least-cost deposits of the exhaustible resource are exploited. To capture competition with oil-based fuels, Coal-To-Liquid becomes competitive (and then enters the market) as soon as oil prices (carbon tax included) exceed a threshold value p_{CTL}. To determine their market potential at a given date, CTL producers form (imperfect) anticipations about future agents’ endogenous decisions in terms of liquid fuel demand *D(t)* and supply by other sources (refined oil and biofuels) *S(t)*. CTL producers are then willing to fill the expected gap by targeting a production level *[D(t)-S(t)]*. But, CTL production may be limited by constraints on delivery capacity due to past

[4]This captures in a simplistic manner the competition between biofuels and oil-based liquid fuels: everything else being equal, the former are more competitive and their penetration into the market is more prominent when higher oil price make the latter more expensive.

[5]We assume that coal is a sufficiently abundant input factor

investment decisions if, due to imperfect foresight, profitability prospects for CTL were underestimated. These prospects are an increasing function of oil prices at each point in time[6] and cumulative investment on CTL over time is then a function of the sum of past oil prices: $p_{cum}(t) = \sum_{i=2010}^{t-\Delta t_{CTL}} p_{oil}(i)$, where the time delay Δt_{CTL} represents investment inertia. The dynamics of investment affects the availability of production capacity and imposes limits on the share s of the targeted CTL production that can be realized at a given date. We adopt a linear dependence between s and cumulative investments measured by $p_{cum}(t)$. As soon as the oil price exceeds p_{CTL}, the contribution of CTL to the supply on liquid fuel markets is given by:

$$CTL(t) = s\left(p_{cum}(t)\right)\left[D(t) - \mathrm{S}(t)\right]. \tag{37}$$

2.2 Electricity generation

The ‘power generation’ Nexus represents an explicit set of 16 standard technologies, either already active or close to maturity.[7] Each of them is characterized by its technico-economic parameters determining the average production discounted cost per kilowatt hour produced. These parameters are: capital costs (dollars per kilowatt installed), energy efficiency (in percentage, for technologies functioning with fossil fuels), exploitation and maintenance costs, fixed or variable costs (respectively in dollars per kilowatt and in dollars per kilowatt hour). The discount rate incorporates capital opportunity cost and a risk factor, which covers both the risk of defect and the social risk associated to controversial technologies (nuclear, CCS). The technico-economic parameters are calibrated either on sectoral technological models (for example the POLES model) or on information from the literature (Grubler et al, 2002; Rao et al, 2006; Sims et al, 2007). They evolve in time according to technical progress, including learning-by doing processes.

Technological choices are based on a minimization of the average production total cost compatible with future electricity demand across six segments of the load curve, representing

[6] Indeed, higher oil prices drive higher prices of liquid fuels, including those produced from coal, and then higher profitability prospects for CTL.

[7] five coal-powered units (Coal Conventional Thermal, Lignite Conventional Thermal, Super Critical Pulverised Coal, Integrated Coal Gasif. Comb. Cycle), two gas-powered units (Gas Conventional Thermal, Gas Turbines Combined Cycle), two oil-powered units (Oil Conventional Thermal, Oil Fired Gas Turbines), two nuclear technologies (standard and new design), three renewables (large hydro, onshore wind, offshore wind). In addition, one technology with CCS is available for coal- and gas-powered units, respectively

the annual fluctuations of electricity demand.[8] This optimal planning procedure for choosing power generation technologies under imperfect anticipations is decomposed into four steps:

- projecting future demand and fuel prices with adaptive anticipations of electricity demand growth over the coming ten years and with future fossil fuels prices.
- choosing renewable production capacities distinguished between hydroelectricity and on-shore and off-shore wind[9], given competition with conventional technologies. The share of each renewable energy in total electricity production is an increasing function of the ratio between its complete production cost per kWh and of the more profitable conventional technology. This share is bounded by the saturation of production potentials and the limits of intermittent production.
- projecting the optimal conventional production park under demand constraint at a 10-year horizon by comparison of unitary discounted production costs among technologies.
- allocating investments to reorient the existing production park towards the ideal anticipated production park by the end of the decade, under the constraints of available capital.

New investment choices affect total production capacity only at the margin, given the inertia in the renewal of the park. We represent the park in capital vintages, and a formerly installed production unit remains available for a certain period in function of its life time. However, available capacities are not necessary mobilized for actual production, which is allocated to production units ensuring lower operational costs. This choice is differentiated along the seven segments of the load curve to represent the different mix of technologies for base and peak production. This assumption allows representing operational flexibility through early retirement of those capacities that, although installed, are not profitable in current economic conditions.

[8] The six segments are divided according to broad categories of annual load length defined by six threshold values between 0h and full year operation (8760h): 8030h, 6570h, 5110h, 3650h, 2190h,. 730h

[9] Wind is the only non-hydraulic renewable energy explicitly represented in the NEXUS. However, solar energy is implicitly represented in “very low energy” .buildings

3. Final energy demand

Historically, the literature on the decoupling between energy and growth has focused on autonomous energy efficiency improvements (implicitly encompassing end-use energy efficiency and structural changes) and on the energy efficiency gap, i.e. the difference between the most energy efficient technologies available and those actually in use.

However important it may be, energy efficiency is not the only driver of energy demand. Indeed, the rate and direction of technical progress and its energy content depend, not only on the transformation of the set of available techniques, but also on the structure of households' demand. This is why the NEXUS endogenize both energy efficiency *stricto sensu,* and the structural change resulting from the interplay between consumption, technology and localization patterns. This enables us to capture the effect of non-energy determinants of energy demand, such as the prices of land and real estate, and political bargaining (set exogenously) over urban infrastructure to be represented. This endogenization of technical change is made for both stationary uses (industry and services, buildings) and non-stationnary uses (freight and passenger transportation).

3.1 Stationnary uses

3.1.1 Industry and services

The industrial and services sectors are represented in an aggregated manner, each of them covering a large variety of economic sub-sectors and products. Technical change then covers not only changes and technical progress in each sub-sector but also the structural effects across sectors. In addition to autonomous energy efficiency gains, the "Industry' and 'services' Nexus represent the structural drop in energy intensity due to a progressive transition from energy-intensive heavy industries to manufacturing industries, and the choice of new techniques which results in both energy efficiency gains and changes in the energy mix.

On the one hand, the progressive switch from industry to services is controlled by saturation levels of per capita consumption of industrial goods (in physical terms, not necessarily in value terms), via an asymptote at κ_{ind} multiplied by its level in 2001. For developing countries, these saturation levels represent various types of catch-up to the consumption style in developed countries.

On the other hand, changes of techniques are driven by operational costs, including energy costs and the other costs linked to their use (capital, maintenance, variable costs). The share of each energy in the new capacities is decided in a logistic function with arguments the total cost of using each energy source and a market heterogeneity parameter measuring the substitutability potentials. In these sectors, these decisions affect the selection of new production capacities but do not influence existing ones. This putty-clay assumption implies that changes in final energy use are dependent on the turnover rate of production capacities, defined by their lifetime Δt_{ind} .

3.1.2 Buildings

The 'Housing and Buildings' nexus represents the dynamics of energy consumption as a function of the energy service level per square meter (heating, cooling, etc.) and the total housing surface.

The former is represented by coefficients encompassing the technical characteristics of the existing stock of end-use equipment and buildings and the increase in demand for energy services: heating, cooking, hot water, lighting, air conditioning, refrigeration and freezing and electrical appliances. The evolution of resulting energy-needs per square meter is captured by coefficients for coal $\alpha_{res}^{coal}(t)$, gas $\alpha_{res}^{gas}(t)$, liquid fuels $\alpha_{res}^{fuel}(t)$ and electricity $\alpha_{res}^{elec}(t)$. These parameters evolve according to the exogenous trajectories calibrated on the outputs of the POLES energy model, which encompass changes in residential energy consumption due to (i) cost variations of the services either due to efficiency gains or energy price variations, (ii) increase in household's income driving access to certain energy services beyond basic needs and (iii) the physical characteristics of buildings (surfaces, insulation, architectural conception).

We also account for the diffusion of "Very Low Energy" buildings at very high energy price, carbon price included. They are represented by a unique alternative housing with annual energy consumption at 50kWh/m2 (80% electricity and 20% gas). The diffusion of this technology in rupture with current trends represents implicitly a multiplicity of advancements, including the autonomous production of energy, the efficient insulation of buildings but also large plans of thermal renovation and regulations reforms in developing countries.

Housing surface per capita has an income elasticity of η_H, and region-specific asymptotes for the floor area per capita, h_{max}. This limit reflects spatial constraints, cultural habits as well as assumptions about future development styles (including the lifestyles in emerging countries vis-à-vis the US, European or Japanese way of life). In the constitution of scenarios, the hypotheses about these asymptotes are made coherent with those concerning the infrastructures of transport, keeping in mind that all are linked to territorial and urban zoning policies.

3.2 Non-Stationnary uses

3.2.1 Freight transport

In the "Transportation NEXUS", the dynamics of the energy intensity of freight transport is driven by an exogenous trend $\mu_f(t)$ and a short-term fuel price elasticity ε_f. They capture autonomous and endogenous energy efficiency gains as well as short-term modal shifts, with the long-term price response resulting from the sequence of those short-term adjustments.

Total energy demand is then driven by freight mobility needs, in turn depending on the level of economic activities and their freight content. Even though the share of transportation in total costs is currently low, decoupling freight mobility demand and economic growth is an important determinant of long-term mitigation costs. In the absence of such a decoupling (constant input-output coefficient), and once efficiency potentials in freight transportation have been exhausted, constraining sectoral carbon emissions from freight transportation would amount to constraining economic activity.

3.2.2 Passenger transport

Passenger mobility needs and their modal breakdown across four travel modes (ground-based public transport, air transport, private vehicles and non-motorized modes) result from the maximization of households' utility under the assumption of constant travel time (Zahavi and Talvitie, 1980) and budget constraints. This helps to represent two crucial determinants of the demand for passenger transportation, namely the induction of mobility demand by infrastructure and the conventional rebound effect consecutive to energy efficiency gains on vehicles (Greening et al, 2000).

The former effect operates through the travel time budget constraint. Indeed, the attractiveness of each transportation mode is determined by vehicle performance and the degree of infrastructure saturation. When mobility demand exceeds the normal load conditions of a given type of infrastructure (e.g., road, airport), speed decreases. In the absence of further investment, households will reallocate their travel time budget to other, more efficient, modes in order to restore efficiency. We can represent the effects of the deployment of alternative infrastructure: a policy in which the building of transportation infrastructure follows the evolution of modal mobility favoring roads for private car mobility *vs.* public policies that redirect part of the investment to railways and other public transport infrastructure.

A drop in mobility costs (mainly the user's car costs), along with progress in the energy efficiency of vehicles, endogenously generates a rebound effect on mobility demand as a result of utility-maximization under income budget constraint. Energy efficiency in private vehicles results from households' decisions on the purchase of new vehicles, based on a mean cost minimization criterion between different types of available technologies (including standard, hybrid and electric vehicles). These vehicles types are differentiated by their capital costs and unitary fuel consumption, the former decreasing in function of the learning-by-doing process at the rate γ for each doubling of cumulated investment in the technology.

In addition to the availability of transportation infrastructure and energy efficiency, mobility needs are dependent upon agents' localization choices (Grazi et al., 2008). This is captured by differences in regional households' motorization rates, everything else being equal (income, energy prices), with dispersed spatial organizations implying a higher dependence on private transport. In each region, the motorization rates increase with disposable per capita income through variable income-elasticity η_{mot}: *(a)* low for very poor people whose access to motorized mobility relies on non-motorized and public modes; *(b)* high for households with a medium per capita income with access to private motorized mobility *(c)* low again, because of saturation effects, for per capita income level comparable to that of the OECD. In addition, the impact of local location choices is represented through basic needs of mobility, which represent the travels imposed by daily journeys (especially, for commuting to work and access to services).

III- Data

1. Calibration

Calibration of the IMACLIM-R model is based on the GTAP database, which provides a set of balanced input-output tables of the world economy. Given the launch date of the RECIPE project, we did not calibrate IMACLIM-R with the more recent GTAP-7 database, but used instead GTAP-6 database (Dimaranan, 2006) which details the world economy in 87 regions and 57 sectors for the year 2001. From this basic material, calibration is done by aggregating the GTAP database according to the IMACLIM-R mapping in 12 regions and 12 sectors and by embarking information from external datasets giving physical quantities for energy and passenger transportation sectors. The elaboration of this hybrid matrix ensuring consistency between money flows and physical quantities is done by modifying input-output tables from the GTAP-6 dataset to make them fully compatible with 2001 energy balances from IEA (in Mtoe) and passenger mobility (in passenger-km) from (Schafer and Victor, 2000). This is done by assuming uniform production prices across uses in each region and for each energy and transport sectors, and substituting money flows reported for those activities in the GTAP-6 database by the expenditures for physical quantities valued at their end-use price, including consumption taxes. This forcing ensures that energy and mobility quantities are preserved, but brings about some adjustments in the input-output tables to restore sectoral supply-use equilibrium conditions in monetary values. This last step is done by reporting the gap in equilibrium conditions in the composite sector.

2. 'Natural Growth' drivers

The natural growth rate of the economy defines the growth rate that the economy would follow if it produced a composite good at full employment, like in standard neoclassical models developed after Solow (1956). It is given by exogenous assumptions on active population and labor productivity growth. Demographic data for active population are derived from (UN, 2005) medium scenario (Table A-3). Labor productivity growth is built upon a convergence hypothesis (Barro and Sala-i-Martin, 1992), the parameters being calibrated on historic trajectories (Maddison, 1995) and 'educated guess' assumptions of long-term trends (Oliveira-Martins et al., 2005). Basically, we assume that USA remains the world leader in productivity per worker with a steady growth of 1.7% per year, whereas the dynamics of productivity in other countries is driven by a partial catch-up. This means that regions with lower absolute productivity per worker in a country experience the faster labor productivity growth (see Table A-4).

Table A-3: *Active population in the IMACLIM-R model (Millions)*

	2001	2010	2030	2050	2100
USA	178	195	203	207	205
Canada	20	22	22	23	20
Europe	374	384	359	330	320
OECD Pacific	132	131	116	100	46
Former Soviet Union	169	178	168	155	126
China	824	930	958	895	827
India	572	702	925	1034	1128
Brazil	104	124	154	167	173
Middle-East	93	122	173	203	250
Africa	397	534	894	1224	1668
Rest of Asia	496	582	706	756	713
Rest of Latin America	193	230	286	309	321

Table A-4: *Average labor productivity growth in the IMACLIM-R model (%)*

	2010-2030	2030-2050	2050-2100
USA	1.9	1.7	1.7
Canada	1.8	1.9	1.7
Europe	2.4	1.9	1.7
OECD Pacific	2.0	1.8	1.7
Former Soviet Union	3.9	2.3	1.7
China	5.8	3.4	1.8
India	5.2	4.2	2.0
Brazil	3.3	2.4	1.7
Middle-East	2.0	2.0	2.0
Africa	2.0	2.0	2.0
Rest of Asia	3.9	3.6	1.8
Rest of Latin America	3.1	2.6	1.8

IV- Regional and sectoral disaggregation

The version of the IMACLIM-R model used in this study divides the world in:

- 12 regions: USA, Canada, Europe, OECD Pacific, Former Soviet Union, China, India, Brazil, Middle-East, Africa, Rest of Asia, Rest of Latin America
- 12 sectors: coal, oil, gas, liquid fuels, electricity, air transport, water transport, other transport, construction, agriculture, energy-intensive industry, services & light-industry.

V- An analytical analysis of the drivers of mitigation costs in second-best economies

We detail here the simplified model used in Section 3.2 of Chapter 3 to identify the major determinants of mitigation costs in the IMACLIM-R model. To this aim, we incorporate the core specificities of second-best macroeconomic interactions in the static equilibrium of the IMACLIM-R model: imperfect competition and imperfect labour markets. To ensure analytical solvability, we consider an economy producing a composite good with energy and labour as input factors.

Imperfect competition is represented through a mark-up pricing rule for the composite good resulting in a margin rate π over production costs:

$$p = p_E e\left(1+\tau_E\right) + wl + \pi p \qquad (38)$$

where e and l are the unitary energy and labour requirements for production, p_E the price of energy, τ_E a tax on energy (taken as a proxy for a carbon tax in case of climate policy) and w the wage rate.[10]

Imperfect labour markets are described by a wage curve introducing an inverse relationship between the real wage rate and unemployment (or under-utilization of the labour force).[11] With Q the total production and L the total labour force, the unemployment rate z is:

[10] Equation (38) is a simplified version of the price equation in IMACLIM-R, which incorporates a term of decreasing static returns when production capacity approaches saturation (see equation (8)).

[11] Microeconometric evidence for such formulation was given in a seminal contribution by (Blanchflower and Oswald 1995) and extensive theories have been developed to support such representation of the labour market (see (Layard et al., 2005), (Lindbeck, 1993) and (Phelps, 1992) for an overview). The basic idea is that high unemployment represents an outside threat that leads workers to accept lower wages as from either the

$$z = 1 - \frac{lQ}{L} \tag{39}$$

The wage curve is then given by:

$$\frac{w}{p} = az^{-\alpha} \tag{40}$$

where, a is a constant and $\alpha > 0$ is the elasticity of the wage curve: the higher α, the more flexible the labour markets.

We introduce Q_0, w_0 and z_0 as the production level, the real wage rate and the unemployment rate in absence of carbon tax. They are implicitly defined by:

$$p_E e\left(1+\tau_E\right) + wl = p_E e + w_0 l \tag{41}$$

$$\frac{p_E e \tau_E}{w_0 l} + \left(\frac{z}{z_0}\right)^{-\alpha} = 1 \tag{42}$$

$$\frac{L}{l} = \frac{Q_0}{1-z_0} \tag{43}$$

Combining equations (38)-(43), the production level Q can then be derived as:

$$Q = \frac{Q_0}{1-z_0} \cdot \left[1 - z_0\left(1 - \frac{p_E . e . \tau_E}{w_0 l}\right)^{-\frac{1}{\alpha}}\right] \tag{44}$$

The variation of activity $\Delta Q = Q - Q_0$ is then given by:

$$\frac{\Delta Q}{Q_0} = \frac{z_0}{1-z_0} \cdot \left[1 - \left(1 - \frac{p_E . e}{w_0 l}\tau_E\right)^{-\frac{1}{\alpha}}\right] \tag{45}$$

This corresponds exactly to equation (1) in Chapter 3.

---

bargaining approach (Layard and Nickell, 1986) or the wage-efficiency approach (Shapiro and Stiglitz, 1984). The former emphasizes the weakening of the power of workers' unions in wage setting negotiations at high unemployment. The latter adopts firms' point of view, who set wages so as to discourage shirking; this level is lower when the threat of not finding a job after being caught shirking gets higher.

References

Amigues JP, Favard P, Gaudet G, and Moreaux M (1998). On the Optimal Order of Natural Resource Use When the Capacity of the Inexhaustible Substitute Is Limited. *Journal of Economic Theory* 80 (1): 153-170

Barro RJ and Sala-i-Martin X (1992). Convergence *Journal of Political Economy* 100(2):223-251

Blanchflower DG, and Oswald AJ (1995). An introduction to the wage curve. *The Journal of Economic Perspectives*: 153-167.

Dimaranan, BV, Editor (2006). *Global Trade, Assistance, and Production: The GTAP 6 Data Base*. *Center for Global Trade Analysis,* Purdue University

Grazi F, van den Bergh JCJM and van Ommeren JN (2008). An Empirical Analysis of Urban Form, Transport, and Global Warming. *The Energy Journal* 29(4), 97-107

Greening L, Greene D and Difiglio C (2000). Energy efficiency and consumption – the rebound effect – a survey. *Energy Policy,* 28(6): 389-401

Grübler A, Nakićenović , Nordhaus W. (Eds.) (2002) Technological Change and the Environment, Washington, DC, Resources for the Future.

Holland S (2003). Extraction Capacity and the Optimal Order of Extraction. *Journal of Environmental Economics and Management* 45(3): 569-588.

IEA (2007). World Energy Outlook 2007. *IEA/OECD*, Paris, France.

Kaufmann R, Dees S, Karadeloglou P and Sanchez M (2004). Does OPEC matter? An econometric analysis of oil prices. *The Energy Journal* 25(4), 67-90.

Kemp MC. and Van Long N (1980). On Two Folk Theorems Concerning the Extraction of Exhaustible Resources. *Econometrica* 48 (3) 663-673

Layard, R., S. Nickell, and R. Jackman. 2005. *Unemployment*. Oxford University Press, Oxford.

Layard, R., and S. Nickell. 1986. Unemployment in Britain. *Economica* 53, no. 210. New Series: S121-S169.

LEPII-EPE (2006). The POLES model, *Institut d'Économie et de Politique de l'Énergie*, Grenoble, France, 12 pp. (*http://webu2.upmfgrenoble.fr/iepe/textes/POLES12pJan06.pdf*).

Lindbeck, A. (1993). *Unemployment and macroeconomics*. MIT Press, Cambridge, MA.

Maddison, A (1995) Monitoring the World Economy: 1820 – 1992 *OECD Development Center, 260 p*

Oliveira Martins J, Gonand F, Antolin P, de la Maisonneuve C and Kwang-Y (2005). The impact of ageing on demand, factor markets and growth *OECD Economics Department Working Papers*, #420, OECD Economics Department

Phelps, ES (1992). Consumer Demand and Equilibrium Unemployment in a Working Model of the Customer-Market Incentive-Wage Economy. *The Quarterly Journal of Economics* 107, no. 3: 1003-1032.

Rao S, Keppo I and Riahi K (2006). Importance of technological change and spillovers in long term climate policy, *The Energy Journal*, (Special issue on Endogenous Technological Change and the Economics of Atmospheric Stabilisation), pp 55–72.

Rehrl T and Friedrich R (2006).. Modeling long-term oil price and extraction with a Hubbert approach: The LOPEX model. *Energy Policy* 34(15):2413-2428.

Reynolds DB (1999). The mineral economy: how prices and costs can falsely signal decreasing scarcity". *Ecological Economics* 31 (1): 155–166.

Schäfer A and Victor DG, (2000). The future mobility of future population .*Transportation Research Part A*, 34: 171-205

Shapiro, C., and J E. Stiglitz. 1984. Equilibrium Unemployment as a Worker Discipline Device. *The American Economic Review* 74, no. 3: 433-444.

Sims REH , Schock RN, Adegbululgbe A, Fenhann J, Konstantinaviciute I, Moomaw W, Nimir HB, Schlamadinger B, Torres-Martínez J, Turner C, Uchiyama Y, Vuori SJV, Wamukonya N, Zhang X (2007). Energy supply, *In Climate Change 2007: Mitigation Contribution of Working Group III to the Fourth Assesent Report of the Intergovernmental Panel on Climate Change* [B Metz, OR Davidson, PR Bosch, R Dave, LA Meyer (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Solow R (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70(1): 65-94.

Uhler RS (1976). Costs and supply in petroleum exploration: the case of Alberta. *Canadian Journal of Economics* 19: 72–90.

Zahavi Y and Talvitie A (1980). Regularities in Travel Time and Money Expenditures. *Transportation Research Record* 750:13-19

Annex B

This appendix provides detailed proofs of the analytical results and discusses the numerical assumptions adopted in Chapter 4.

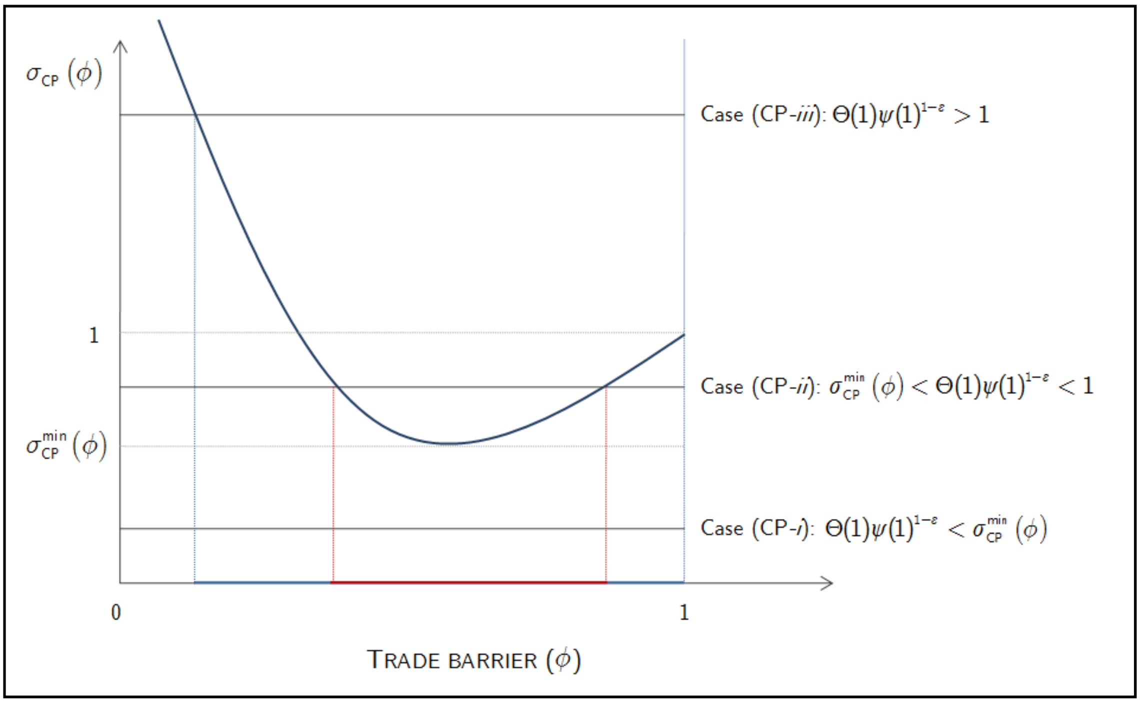
I. Long-Run Equilibrium Conditions for Symmetric Spatial Configurations

Proof of Proposition 1: From (25) it follows that: $\frac{\partial \sigma_{\text{CP}}}{\partial \phi} = \frac{\left(1+\frac{\delta}{\varepsilon}\right)\left(1+\frac{\delta}{\varepsilon-1}\right)\left(\phi^2-\phi_m^2\right)}{2\phi^{2+\frac{\delta}{\varepsilon-1}}}$; with: $\phi_m = \sqrt{\frac{(\varepsilon-\delta)(\varepsilon-1-\delta)}{(\varepsilon+\delta)(\varepsilon-1+\delta)}}$. The value $\frac{\partial \sigma_{\text{CP}}}{\partial \phi}$ is then negative for $\phi < \phi_m$, while it is positive for $\phi > \phi_m$. This leads to σ_{CP} being decreasing for $\phi < \phi_m$ while increasing for $\phi > \phi_m$. The σ_{CP} function reaches its minimum at $\phi = \phi_m$, and the minimum value $\sigma_{\text{CP}}^{\min}$ of $\sigma_{\text{CP}}(\phi)$ is given by $\sigma_{\text{CP}}(\phi_m)$:

$$\sigma_{\text{CP}}^{\min} = \frac{\varepsilon-1}{\varepsilon}\left(\frac{\varepsilon-\delta}{\varepsilon-1+\delta}\right)^{\frac{1}{2}\left(\frac{\delta}{\varepsilon-1}+1\right)}\left(\frac{\varepsilon-1-\delta}{\varepsilon+\delta}\right)^{\frac{1}{2}\left(\frac{\delta}{\varepsilon-1}-1\right)}.$$

In addition, $\lim_{\phi\to 0}\sigma_{\text{CP}}(\phi) = +\infty$ and $\sigma_{\text{CP}}(\phi=1)=1$. Figure B.1 gives a graphical illustration of $\sigma_{\text{CP}}(\phi)$.

Figure B-1: *Graphical Illustration of the Core-periphery Pattern: Three Cases.*



The stability pattern of the core-periphery depends on the value of $\Theta(1)\psi(1)^{1-\varepsilon}$:

- If $\Theta(1)\psi(1)^{1-\varepsilon} < \sigma_{\text{CP}}^{\min}$, then we have $\Theta(1)\psi(1)^{1-\varepsilon} < \sigma_{\text{CP}}(\phi)$ for any ϕ. According to Condition 1, this means that the core-periphery is never an equilibrium. This corresponds to case CP–*i* in Proposition 1;
- If $\sigma_{\text{CP}}^{\min} < \Theta(1)\psi(1)^{1-\varepsilon} < 1$, then there exist two ϕ-values over $\phi \in [0;1]$ such that $\Theta(1)\psi(1)^{1-\varepsilon} = \sigma_{\text{CP}}(\phi)$. By noting these $\underline{\phi}_s$ and $\overline{\phi}_s$, the condition $\Theta(1)\psi(1)^{1-\varepsilon} > \sigma_{\text{CP}}(\phi)$ is satisfied if and only if $\phi \in \left[\underline{\phi}_s;\overline{\phi}_s\right]$. According to Condition 1, this means that the core-periphery is an equilibrium if and only if $\phi \in \left[\underline{\phi}_s;\overline{\phi}_s\right]$. This corresponds to case CP–*ii* in Proposition 1;
- If $\Theta(1)\psi(1)^{1-\varepsilon} > 1$, then there exist only one ϕ-value in the interval $\phi \in [0;1]$ such that $\Theta(1)\psi(1)^{1-\varepsilon} = \sigma_{\text{CP}}(\phi)$. By noting this value $\underline{\phi}_s$, the condition $\Theta(1)\psi(1)^{1-\varepsilon} > \sigma_{\text{CP}}(\phi)$ is satisfied if and only if $\phi \in \left[\underline{\phi}_s;1\right]$. According to Condition 1, this means that the core-periphery is an equilibrium if and only if $\phi \in \left[\underline{\phi}_s;1\right]$. This corresponds to case CP–*iii* in Proposition 1.

This concludes the proof.

Proof of Proposition 2: From equation (28), σ_{SS} can be rewritten as a second-order polynomial in ϕ, as follows: $\sigma_{\text{SS}}(\phi) = a_0 + a_1\phi + a_2\phi^2$. Coefficients a_0, a_1, a_2 can be expressed in terms of the constants of the model δ, ε, a^f, m^f, L and the intensity of the agglomeration and environmental effects $d_\psi^{(h=0.5)}$ and $d_\Theta^{(h=0.5)}$:

$$a_0 = \frac{(\varepsilon-\delta)\left[d_\psi^{(h=0.5)}\delta(\varepsilon-1) - 4(\varepsilon-1-\delta)\right]}{2(\varepsilon-1)};$$

$$a_1 = \frac{4\delta^2 + 4\varepsilon(\varepsilon-1)}{\varepsilon-1} + d_\psi^{(h=0.5)}\left[\delta^2 + 2\varepsilon(\varepsilon-1)\right] - d_\Theta^{(h=0.5)}\frac{\delta a^f L\varepsilon(\varepsilon-1)}{2\gamma(\varepsilon-\delta)}\left(4 + d_\psi^{(h=0.5)}(\varepsilon-1)\right);$$

$$a_2 = -\frac{(\varepsilon+\delta)\left[d_\psi^{(h=0.5)}\delta(\varepsilon-1) + 4(\varepsilon-1+\delta)\right]}{2(\varepsilon-1)}.$$

The coefficient a_2 of the polynomial $\sigma_{\text{SS}}(\phi)$ is negative, so that $\sigma_{\text{SS}}(\phi)$ is a concave function. The sign of $\sigma_{\text{SS}}(\phi)$ over $\phi \in [0;1]$ is dependent on its signs at $\phi = 0$ and $\phi = 1$. Four cases must be distinguished:

1) If $\sigma_{SS}(0)>0$ and $\sigma_{SS}(1)>0$, the concave polynomial $\sigma_{SS}(\phi)$ remains positive over the whole range $\phi\in[0;1]$;

2) If $\sigma_{SS}(0)<0$ and $\sigma_{SS}(1)>0$, the polynomial $\sigma_{SS}(\phi)$ has a single root ϕ_b over $\phi\in[0;1]$; $\sigma_{SS}(\phi)$ is negative over $\phi\in[0;\phi_b]$, while positive over $\phi\in[\phi_b;1]$;

3) If $\sigma_{SS}(0)>0$ and $\sigma_{SS}(1)<0$, the polynomial $\sigma_{SS}(\phi)$ has a single root ϕ_b over $\phi\in[0;1]$; $\sigma_{SS}(\phi)$ is positive over $\phi\in[0;\phi_b]$, while negative over $\phi\in[\phi_b;1]$;

4) If $\sigma_{SS}(0)<0$ and $\sigma_{SS}(1)<0$, the polynomial $\sigma_{SS}(\phi)$ has either zero or two roots according to the sign of the discriminant $\Delta=a_1^2-4a_0a_2$

   a) If $\Delta>0$, $\sigma_{SS}(\phi)$ has two roots $\underline{\phi}_b$ and $\overline{\phi}_b$ ($\underline{\phi}_b<\overline{\phi}_b$); $\sigma_{SS}(\phi)$ is positive $\phi\in\left[\underline{\phi}_b;\overline{\phi}_b\right]$, while negative for $\phi<\underline{\phi}_b$ and $\phi>\overline{\phi}_b$. It can be demonstrated that for $0<\underline{\phi}_b<\overline{\phi}_b<1$:

      i. $\Delta>0$ leads to $\dfrac{\partial\sigma_{SS}}{\partial\phi}(0)>0$. With $\sigma_{SS}(0)<0$, this means that $\underline{\phi}_b>0$;

      ii. $\sigma_{SS}(1)<0$ leads to $\dfrac{\partial\sigma_{SS}}{\partial\phi}(1)<0$. With $\sigma_{SS}(1)<0$, this means $\overline{\phi}_b<1$.

      Then, $\sigma_{SS}(\phi)$ is positive over $\phi\in\left[\underline{\phi}_b;\overline{\phi}_b\right]$, while negative over $\phi\in[0;1]\backslash\left[\underline{\phi}_b;\overline{\phi}_b\right]$;

   b) If $\Delta<0$, $\sigma_{SS}(\phi)$ has no root and remains negative over the whole range $\phi\in[0;1]$.

   c) From eq. (25) and (26) it follows that:

$$\sigma_{SS}(0)>0\Leftrightarrow d_\psi^{(0.5)}>d_{\psi,0};$$
$$\sigma_{SS}(1)>0\Leftrightarrow d_\Theta^{(0.5)}<\zeta\left(d_\psi^{(0.5)}\right);$$
$$\Delta>0\Leftrightarrow d_\Theta^{(0.5)}<\zeta_\Delta\left(d_\psi^{(0.5)}\right),$$

where the symbols denote the following mathematical expressions:

$$d_{\psi,0} = \frac{4(\varepsilon - 1 - \delta)}{\delta(\varepsilon - 1)};$$

$$\zeta\left(d_{\psi}^{(0.5)}\right) = \frac{4\gamma(\varepsilon - \delta)}{\delta a^f L} \frac{(\varepsilon - \delta)}{4 + d_{\psi}^{(0.5)}(\varepsilon - 1)};$$

$$\zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right) = \zeta\left(d_{\psi}^{(0.5)}\right) + \frac{2\gamma(\varepsilon - \delta)\left\{4(\varepsilon - 1) + \varepsilon\left[4 + d_{\psi}^{(0.5)}(\varepsilon - 1)\right]\right\}^2}{\varepsilon(\varepsilon - 1)^2 \delta a^f L\left[4 + d_{\psi}^{(0.5)}(\varepsilon - 1)\right]} \cdot$$

$$\cdot \frac{\delta^2}{4\delta^2 + 4\varepsilon(\varepsilon - 1) + \delta^2 d_{\psi}^{(0.5)}(\varepsilon - 1) + \sqrt{\left(\varepsilon^2 - \delta^2\right)\left\{\left[4(\varepsilon - 1)\right]^2 - \delta^2\left[4 + d_{\psi}^{(0.5)}(\varepsilon - 1)\right]^2\right\}}}.$$

The above conditions (1 to 4) can then be re-stated in a simpler way, as follows:

1) If $d_{\psi}^{(0.5)} > d_{\psi,0}$ and $d_{\Theta}^{(0.5)} < \zeta\left(d_{\psi}^{(0.5)}\right)$, $\sigma_{\text{SS}}(\phi)$ is positive over the whole range $\phi \in [0;1]$. With Condition 2, this corresponds to case SS–*i* in Proposition 2;
2) If $d_{\psi}^{(0.5)} < d_{\psi,0}$ and $d_{\Theta}^{(0.5)} < \zeta\left(d_{\psi}^{(0.5)}\right)$, $\sigma_{\text{SS}}(\phi)$ is negative over $\phi \in [0;\phi_b]$ and positive over $\phi \in [\phi_b;1]$. With Condition 2, this corresponds to case SS–*ii* in Proposition 2;
3) If $d_{\psi}^{(0.5)} > d_{\psi,0}$ and $d_{\Theta}^{(0.5)} > \zeta\left(d_{\psi}^{(0.5)}\right)$, $\sigma_{\text{SS}}(\phi)$ is positive over $\phi \in [0;\phi_b]$, and negative over $\phi \in [\phi_b;1]$. With Condition 2, this corresponds to case SS–*iii* in Proposition 2;
4) If $d_{\psi}^{(0.5)} < d_{\psi,0}$, and $d_{\Theta}^{(0.5)} > \zeta\left(d_{\psi}^{(0.5)}\right)$:
    a) If $d_{\Theta}^{(0.5)} < \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, then $\sigma_{\text{SS}}(\phi)$ is positive over $\phi \in \left[\underline{\phi}_b;\overline{\phi}_b\right]$, and negative over $\phi \in [0;1] \setminus \left[\underline{\phi}_b;\overline{\phi}_b\right]$. With Condition 2, this corresponds to case SS–*iv* in Proposition 2;
    b) If $d_{\Theta}^{(0.5)} > \zeta_{\Delta}\left(d_{\psi}^{(0.5)}\right)$, then $\sigma_{\text{SS}}(\phi)$ remains negative over the whole range $\phi \in [0;1]$. With Condition 2, this corresponds to case SS–*v* in Proposition 2.

This concludes the proof.

II. Long-Run Partial Agglomeration in Symmetric Spatial Configurations

Proof of Proposition 3: For a given ϕ satisfying $0 \le \phi \le 1$, we define a function g_ϕ such that: $\forall h \in [0;1], g_\phi(h) = \Omega(h,\phi)$. Let us consider case (PA– *i*). This means that the ϕ-value is chosen so that:

- The core-periphery $h = 1$ is a stable equilibrium. According to (19-*b*), this means that: $g_\phi(1) > 0$;
- The symmetric spreading $h = 0.5$ is a stable equilibrium. According to (19-*a*), this means that: $g_\phi(0.5) = 0$ and $g'_\phi(0.5) < 0$.

By continuity of function g_ϕ the last two conditions mean that there exists a value $\overline{h}$ such that $0.5 < \overline{h} < 1$ and $g_\phi\left(\overline{h}\right) < 0$. Conditions $g_\phi\left(\overline{h}\right) < 0$ and $g_\phi(1) > 0$ mean that there exists a value h_0 such that:

$$\begin{Bmatrix} \overline{h} < h_0 < 1 \\ g_\phi\left(h_0\right) = 0 \\ g'_\phi\left(h_0\right) > 0 \end{Bmatrix}.$$

This value h_0 is then an unstable partial agglomeration equilibrium. The proof is similar for case (PA– *ii*). This concludes the proof.

III. Long-Run Partial Agglomeration in Non-Symmetric Spatial Configurations

Proof of Proposition 4: For a given ϕ satisfying $0 \le \phi \le 1$, we define a function g_ϕ such that: $\forall h \in [0;1], g_\phi(h) = \Omega(h,\phi)$. Let us consider case (PA*'* –*i*). This means that the ϕ-value is chosen so that:

- The core-periphery $h = 1$ is a stable equilibrium. According to (19-*b*), this means that: $g_\phi(1) > 0$;
- The core-periphery $h = 0$ is a stable equilibrium. According to (19-*c*), this means that: $g_\phi(0) < 0$.

By continuity of function g_ϕ, this mean that there exists a value h_0 such that:

$$\begin{Bmatrix} 0 < h_0 < 1 \\ g_\phi\left(h_0\right) = 0 \\ g'_\phi\left(h_0\right) > 0 \end{Bmatrix}.$$

This value h_0 is an unstable partial agglomeration equilibrium. The proof is similar for case (PA*'–ii*). This concludes the proof.

IV. Dependence of Environmental and Agglomeration Effects on Trade Barriers

Here, we demonstrate the following three results regarding the dependence of agglomeration and environmental effects on the value of trade barriers ϕ:

The agglomeration effect is more intense at low trade barriers: This comes down to demonstrating that $\sigma_{\mathrm{SS}}^{(\psi)}\left(\phi\right)$ in (28) is increasing in ϕ. The function $\sigma_{\mathrm{SS}}^{(\psi)}\left(\phi\right)$ is a second-order polynomial in ϕ. Its dominant term $-\frac{1}{2}\delta\left(\varepsilon+\delta\right)$ is negative, so that $\sigma_{\mathrm{SS}}^{(\psi)}\left(\phi\right)$ is a concave function. It reaches a minimum at $\phi = \phi_m^{(\psi)}$, implicitly defined by $\frac{\partial \sigma_{\mathrm{SS}}^{(\psi)}}{\partial \phi}\left(\phi_m^{(\psi)}\right) = 0$. Function $\sigma_{\mathrm{SS}}^{(\psi)}\left(\phi\right)$ is increasing for $\phi < \phi_m^{(\psi)}$, while decreasing for $\phi > \phi_m^{(\psi)}$. From equation (28), it follows that $\phi_m^{(\psi)} = \frac{4\varepsilon\left(\varepsilon-1\right)+2\delta^2}{2\delta\left(\varepsilon+\delta\right)}$. With the "no black hole" condition $\varepsilon - 1 > \delta$, we have $\frac{4\varepsilon\left(\varepsilon-1\right)+2\delta^2}{2\delta\left(\varepsilon+\delta\right)} > \frac{\varepsilon}{\varepsilon+\delta}+1$, so that $\phi_m^{(\psi)} > 1$. This means in particular that $\sigma_{\mathrm{SS}}^{(\psi)}\left(\phi\right)$ is increasing over $\phi \in \left[0;1\right]$. This demonstrates the result.

The environmental effect is more intense at low trade barriers: This comes down to demonstrating that function $\sigma_{\mathrm{SS}}^{(\Theta)}\left(\phi\right)$ in (28) is increasing in ϕ. The function $\sigma_{\mathrm{SS}}^{(\Theta)}\left(\phi\right)$ is linear in ϕ, with a positive multiplicative term $\frac{\delta\varepsilon a^f L\left(\varepsilon-1\right)}{\gamma\left(\varepsilon-\delta\right)}\left(2+d_\psi^{(0.5)}\frac{\varepsilon-1}{2}\right)$, so that $\sigma_{\mathrm{SS}}^{(\Theta)}\left(\phi\right)$ is uniformly increasing. This demonstrates the result.

The environmental effect is stronger than the agglomeration effect at low trade barriers: This comes down to demonstrating that the ratio $\dfrac{\sigma_{\mathrm{SS}}^{(\Theta)}(\phi)}{\sigma_{\mathrm{SS}}^{(\psi)}(\phi)}$ is increasing in ϕ. From equation (28) it follows that the ratio $\dfrac{\sigma_{\mathrm{SS}}^{(\Theta)}(\phi)}{\sigma_{\mathrm{SS}}^{(\psi)}(\phi)} = \dfrac{\dfrac{\delta\varepsilon a^f L(\varepsilon-1)}{\gamma(\varepsilon-\delta)}\left(4+d_{\psi}^{(0.5)}(\varepsilon-1)\right)}{\left[4\varepsilon(\varepsilon-1)+2\delta^2\right]+\dfrac{1}{\phi}\delta(\varepsilon-\delta)-\delta\phi(\varepsilon+\delta)}$. Since the denominator is a decreasing function in ϕ, $\dfrac{\sigma_{\mathrm{SS}}^{(\Theta)}(\phi)}{\sigma_{\mathrm{SS}}^{(\psi)}(\phi)}$ is increasing in ϕ.

V. Values of Model Parameters and Exogenous Variables

Parameters and exogenous variables defining the economy: In line with the literature [Grazi, van den Bergh and Rietveld (2007)], the exogenous variable total unskilled labor availability L is set equal to 5. We normalize the global skilled population to 1, i.e. $H = H_1 + H_2 = 1$. Whenever possible, the values of the economic parameters have been taken from the literature on spatial and trade economics [e.g., Fujita, Krugman and Venables (1999); Fujita and Thisse (2002); Bernard, Eaton, Jensen and Kortum (2003)]. The share of income spent on manufactured goods in eq. (1) is set equal $\delta = 0.4$. The elasticity of substitution in eq. (1) is $\varepsilon = 3$. Finally we assume a one-to-one production structure in the energy sector (one unskilled worker produces one unit of energy), which comes down to setting the labor requirement parameter in (6) $\gamma = 1$.

Parameters defining local pollution E_j^L and global pollution E_j^G: Concerning the local pollution parameters, the parameter a^L in eq. (13) is normalized to 1, as a definition of the unit of measure of pollution-externality flow arising from manufacturing production. As for the parameters defining the global pollution E_j^G in eq. (27), a^G is normalized to 1. The numerical value of parameter b^G is calibrated on empirical data for manufacturing production and trade. Based on estimates by the World Trade Organization and the World Bank, respectively, trade of manufacturing goods τ_{tot}^M amounted to 8, 257 Billion$ in 2006, and total industrial production x^M_{tot} to 13,428 Billion$.[120] These two activities represent, respectively, 3.5% and 25% of world greenhouse gas (GHG) emissions from energy use considered as the stock pollutant under consideration [International Energy Agency (IEA) (2008)]. According to (27), the amount of emissions associated with production x^M_{tot} and trade τ^M_{tot} are given by $a^G x^M_{tot}$ and $b^G \tau^M_{tot}$, respectively. The above figures on relative GHG emissions then lead to $\frac{b^G \tau^M_{tot}}{a^G x^M_{tot}} = \frac{3.5\%}{25\%} = 0.14$. This gives $b^G = 0.23$.

The market-form parameter β_j: The parameter β_j captures the exogenous spatial characteristics of region *j* in terms of the degree of (transport and electricity) infrastructure development characterizing the economy's configuration (see Section II.A). Regional spatial structure alters the energy intensity of production activities that are located in *j*. Two types of regional spatial structure are considered: one is characterized by an 'urbanized' region, with a high degree of infrastructure development (as captured by a low value of β_j); another by a less urbanized, 'undeveloped' region, with little land development (high value of β_j). When considering configurations with *symmetric* spatial structure, the β_j parameters enter the indirect utility differential only through the multiplicative term $\frac{\Gamma'}{\beta_j^{\delta}}$ (see eq. (21) and eq. (22) in the case $\beta_1 = \beta_2$). This term is constant and strictly positive and, hence, a change in the *β*-value does not modify the stability conditions in (19). When *asymmetric* configurations are considered, the *β*-parameters enter the indirect utility differential through the ratio $\left(\frac{\beta_2}{\beta_1}\right)^{1-\varepsilon}$ (see eq. (19) and eq. (20) in the case $\beta_1 \neq \beta_2$). In this case, long-run development patterns driven by the utility differential depend entirely on the relative numerical values of the parameters. Without loss of generality, we set $\beta_1 = 1$ and let the numerical value of the ratio $\frac{\beta_2}{\beta_1}$ be calibrated over some alternative trend of energy-intensity of the domestic economy between two comparable regions, such as, e.g. the USA and Europe. For the year 2006, official data from the EIA give an energy intensity of economic activity (amount of energy

[120] Data from the International Trade Statistics 2007, published by the World Trade Organization, available at http://www.wto.org/english/res_e/statis_e/its2007_e/its07_toc_e.htm.

used per unit of value added) of 8840 Btu/$ in the USA and 6536 Btu/$ in Europe.[121] The β-parameters capture these differences in energy intensity, so that we obtain $\frac{\beta_2}{\beta_1} = \frac{8840}{6536} \approx 1.35$.

With $\beta_1 = 1$, this leads $\beta_2 = 1.35$ and $\nu = \left(\frac{\beta_2}{\beta_1}\right)^{1-\varepsilon} \approx (1.35)^{1-\varepsilon} = 0.55$.

Table B-1: *Values of the agglomeration parameters in the spatial configurations*

Spatial Configuration	β_1 **(Region 1)**	β_2 **(Region 2)**
A (both regions with undeveloped land)	1.35	1.35
B (both regions with urbanized land)	1	1
C (one region urbanized, other undeveloped)	1	1.35

VI. Functional Specifications

The market-density effect function $\bar{\psi}(n_j)$: Function $\bar{\psi}(n_j)$ captures the decrease of energy-related production costs resulting from agglomeration of firms in the *j* region. By assumption, this function is decreasing in n_j and satisfies condition: $\bar{\psi}(0) = 1$. Given the relation between the number of active firms in region *j* and the amount of skilled workers regionally employed (see eq.(9)) and defining $h = \frac{H_1}{H}$ as the share of the regional population, the 'market size effect' $\bar{\psi}(n_j)$ can be re-written as a function of *h*: $\psi(h)$. We choose to adopt an exponential mathematical form: $\psi(h) = e^{-\mu_\psi h}$ where μ_ψ is a positive constant. Such a function satisfies $\psi''(h) > 0$ so that the agglomeration effect features decreasing returns to agglomeration: production costs are less reduced by a marginal increase in the degree of agglomeration if production is already intensely agglomerated. This exponential mathematical form is convenient since it leads to simple analytical expressions for $d_\psi^{(0.5)}$ providing a straightforward interpretation of parameter μ_ψ. Indeed, since $d_\psi^{(0.5)} = 2\mu_\psi$ it follows that μ_ψ measures directly the intensity of the market size effect, ψ. We choose numerical values of

[121] See http://www.eia.doe.gov/emeu/international/energyconsumption.html.

μ_ψ that allow for considering non-trivial cases in which the effect of agglomeration is strong enough to affect agents' location choices. This comes down to assuming a strong agglomeration effect, which corresponds to the analytical condition $d_\psi^{(0.5)} > d_{\psi,0}$ (see section II-A), with $d_\psi^{(0.5)} = 2\mu_\psi$. This can be rewritten as $\mu_\psi > \frac{d_{\psi,0}}{2}$. Taking $\delta = 0.4$ and $\varepsilon = 3$, and recalling $d_{\psi,0} = \frac{4(\varepsilon - 1 - \delta)}{\delta(\varepsilon - 1)}$ gives $d_{\psi,0} = 8$ and hence $\mu_\psi > 4$. For the ease of computation, we take $\mu_\psi = 5$. We are then able to study cases where positive agglomeration effects play an important role. Such cases have never been investigated in the literature because of the inherent limitations of existing frameworks in which the agglomeration effect is not measurable.

The environmental-externality function $\Theta\left(E_j^L\right)$: It captures the decrease of utility due to the effect of negative local environmental externalities. We posit this function to be decreasing and satisfy condition $\Theta(0) = 1$. We set $\Theta\left(E_j^L\right) = 2 - e^{\mu_\Theta E_j^L}$ where μ_Θ is a positive constant capturing the intensity of the negative environmental effect. This function satisfies the condition: $\Theta''\left(E_j^L\right) = -\left(\mu_\Theta\right)^2 e^{\mu_\Theta E_j^L} < 0$, to capture the non-linear response of environmental damage to pollution. We choose numerical values of μ_Θ that allow consideration of non-trivial cases in which the environmental effect is not fully dominating the agglomeration and trade effects, as this case is of little relevance to a thorough analysis of sustainability, in which agglomeration and trade do matter. Recalling that $d_\Theta^{(0.5)} > \zeta\left(d_\psi^{(0.5)}\right)$ corresponds to a 'strong' environmental effect, whereas $d_\Theta^{(0.5)} < \zeta\left(d_\psi^{(0.5)}\right)$ captures a 'weak' environmental effect (see section II.1), we retain the case $d_\Theta^{(0.5)} = \zeta\left(d_\psi^{(0.5)}\right)$, which corresponds to the environmental effect taking a moderate, mean intensity on utility. This analytical condition leads to setting a numerical value of $\mu_\Theta = 0.45$.

Annex C

This appendix provides a list of all 74 agglomerations considered in Chapter 5 and displays the details of the results obtained for all of them.

I-List of agglomerations

USA: Atlanta, Baltimore, Boston, Chicago, Cleveland, Dallas, Denver, Detroit, Houston, Los Angeles, Miami, Minneapolis, New York, Philadelphia, Phoenix, Pittsburgh, Portland, San Diego, San Francisco, Seattle, St.Louis, Tampa Bay, Washington.

Canada: Montreal, Toronto, Vancouver.

Europe: Ankara, Athens, Barcelona, Berlin, Birmingham, Brussels, Budapest, Copenhagen, Dublin, Frankfurt, Hanburg, Helsinki, Istanbul, Izmir, Krakow, Leeds, Lille, Lisbon, London, Lyon, Madrid, Manchester, Milan, Munich, Naples, Oslo, Paris, Prague, Rand-Holland, Rhine-Ruhr, Rome, Stockholm, Stuttgart, Turin, Valencia, Vienna, Warsaw, Zurich.

OECD Pacific: Aichi, Auckland, Busan, Deagu, Fukuoka, Melbourne, Osaka, Seoul, Sydney, Tokyo.

II- Data for calibration

Table C-1*: Numerical value of calibration variables at the agglomeration level*

EUROPE						USA					
	$L_j^{(0)}$	$d_j^{(0)}$	$Q_j^{(0)}$	$w_j^{(0)}$	$CC_j^{(0)}$		$L_j^{(0)}$	$d_j^{(0)}$	$Q_j^{(0)}$	$w_j^{(0)}$	$CC_j^{(0)}$
Units	*thousands*	*km*	*Index (1=Paris)*	*Index (1=Paris)*	*%*	*Units*	*thousands*	*km*	*Index (1=New York)*	*Index (1=New York)*	*%*
Rhine-Ruhr	5894	72	0.81	1.14	15	New York	8285	54	1.00	1.00	11
Paris	5510	62	1.00	1.00	15	Los Angeles	6265	43	0.56	0.69	11
Istanbul	4730	55	0.25	0.43	13	Chicago	4390	55	0.44	0.72	11
Rand-Holland	3896	51	0.54	1.05	15	Philadelphia	2775	39	0.27	0.68	11
London	3624	22	0.65	1.17	15	Miami	2550	44	0.20	0.58	11
Milan	3360	63	0.59	0.88	15	Washington	2426	44	0.29	0.81	11
Berlin	2759	98	0.29	1.07	15	Atlanta	2418	52	0.23	0.69	11
Munich	2723	99	0.48	1.14	15	Dallas	2393	54	0.28	0.72	11
Madrid	2560	51	0.36	1.02	15	San Francisco	2287	33	0.26	1.01	11
Frankfurt	2491	73	0.41	1.14	15	Boston	2232	38	0.26	0.77	11
Barcelona	2317	50	0.29	0.99	15	Houston	2186	57	0.26	0.72	11
Hanburg	2023	84	0.32	1.14	15	Detroit	2146	37	0.20	0.72	11
Athens	1645	35	0.22	0.74	15	Phoenix	1720	68	0.14	0.60	11
Rome	1621	41	0.28	0.69	15	Minneapolis	1547	45	0.16	0.69	11
Brussels	1620	45	0.29	1.05	15	Seattle	1478	45	0.17	0.77	11
Ankara	1575	88	0.07	0.37	13	San Diego	1409	38	0.13	0.65	11
Izmir	1462	62	0.07	0.39	13	St.Louis	1352	55	0.11	0.61	11
Zurich	1395	39	0.19	1.22	15	Baltimore	1274	31	0.11	0.63	11
Lisbon	1379	31	0.15	0.89	15	Denver	1245	52	0.12	0.72	11
Warsaw	1287	49	0.15	0.65	13	Tampa Bay	1218	33	0.09	0.54	11
Copenhagen	1286	54	0.16	1.00	15	Pittsburgh	1197	41	0.10	0.59	11
Budapest	1231	47	0.13	0.61	13	Cleveland	1047	36	0.09	0.60	11
Stuttgart	1223	34	0.21	1.14	15	Portland	912	47	0.09	0.55	11
Manchester	1201	20	0.14	0.85	15	**CANADA**					
Prague	1200	60	0.12	0.58	13	Toronto	2709	43	1.00	1.00	11
Birmingham	1169	17	0.15	0.90	15						
Stockholm	1165	69	0.18	0.95	15	Montreal	1854	37	0.58	0.82	11
Vienna	1083	38	0.18	1.10	13	Vancouver	1113	30	0.36	0.84	11
Naples	1079	19	0.12	0.58	15	**OECD PACIFIC**					
Lille	1067	43	0.12	1.00	15	Tokyo	18238	65	1.00	1.00	15
Valencia	1034	59	0.10	0.96	15	Seoul	10555	61	0.37	0.90	15
Leeds	1019	25	0.12	0.84	15	Osaka	8695	68	0.42	0.88	15
Krakow	987	50	0.05	0.52	13	Aichi	4926	58	0.25	0.94	15
Turin	978	46	0.15	0.76	15	Busan	3635	63	0.14	0.95	15
Helsinki	965	79	0.13	1.02	15	Fukuoka	2568	39	0.11	0.79	15
Oslo	918	89	0.17	0.87	15	Sydney	2168	62	0.12	1.19	15
Dublin	773	47	0.12	1.24	15	Melbourne	1833	50	0.10	1.14	15
Lyon	717	32	0.12	1.00	15	Deagu	1177	17	0.03	0.58	15
Lyon	717	32	0.12	1.00	15	Auckland	602	38	0.03	0.82	15

Table C-2: Numerical value of macroeconomic aggregates in 2001 for the four macro-regions

Macroeconomic aggregates in 2001	**Units**	**USA**	**CAN**	**EUR**	**OECD PACIFIC**
Value Added $\overline{V^{(0)}}$	**10^{12}\$**	*18.2*	*1.3*	*17.1*	*9.1*
Effective labor $\overline{S^{(0)}}$	**10^{6}worker**	*142.6*	*16.0*	*299*	*105.5*
Wage rate $\overline{w^{(0)}}$	**10^{3}\$/worker**	*45.5*	*23.3*	*14.8*	*25.0*
Population $\overline{Pop^{(0)}}$	**10^{6} persons**	*285.3*	*31.1*	*588.2*	*204.7*

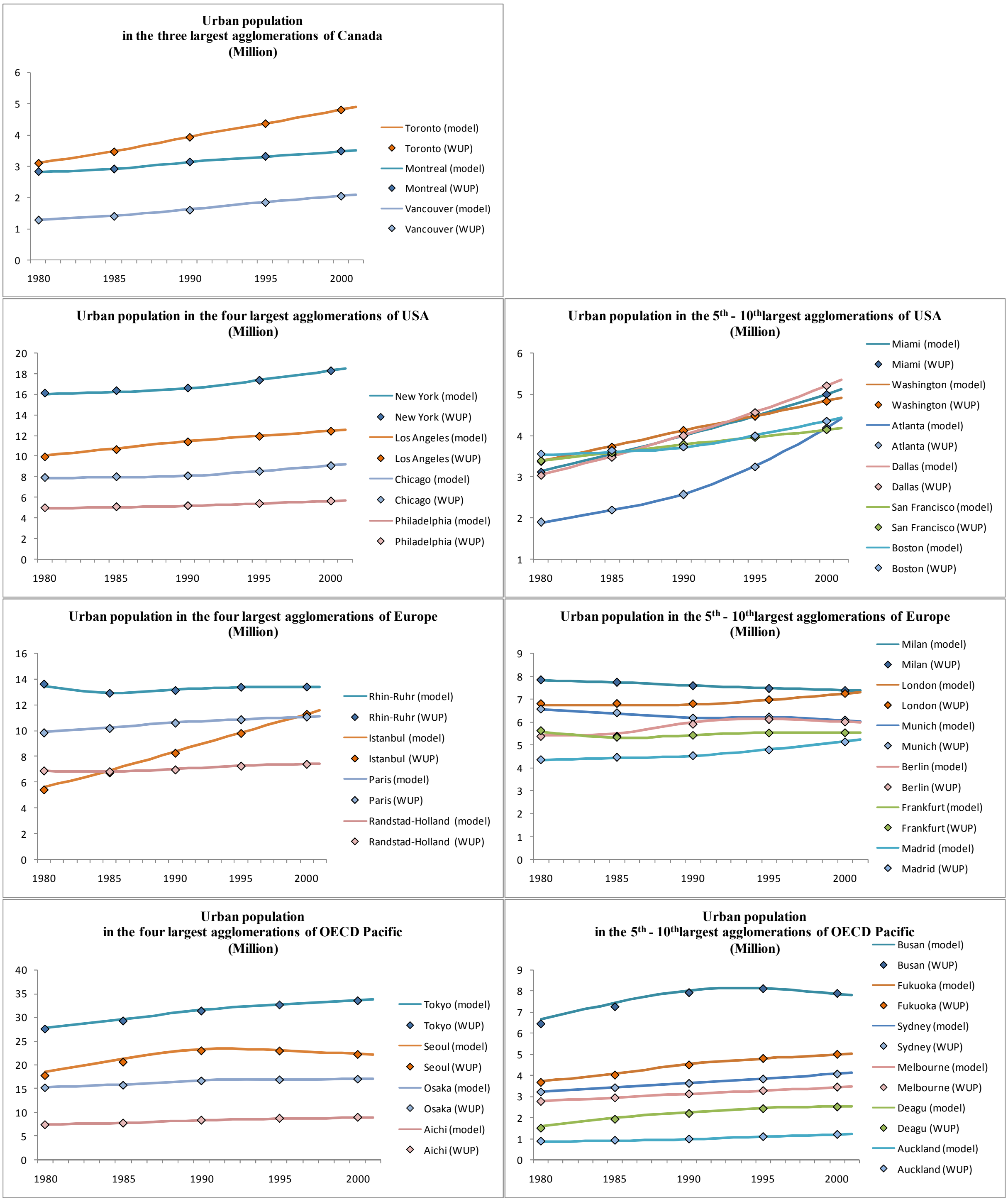
III-Validation on past trends

Table C-3: Average growth rate of macroeconomic variables over the period 1980-2001.

Average annual growth rate of macroeconomic aggregates (%)							
Macro economic aggregate	**Variable name**	**Region**	**Time period**				**Data source**
			1980-1985	**1985-1990**	**1990-1995**	**1995-2001**	
Value added V(t)	g_V	**USA**	*2.5*	*3.7*	*2.6*	*3.9*	*a*
		Canada	*2.3*	*3.4*	*1.5*	*3.7*	*d*
		Europe	*1.9*	*2.6*	*1.4*	*2.8*	*d*
		JANZ	*3.2*	*4.4*	*2.4*	*0.9*	*d*
Labor productivity l(t)	g_l	**USA**	*1.4*	*1.8*	*1.8*	*2.2*	*b*
		Canada	*1.1*	*0.7*	*1.2*	*1.9*	*d*
		Europe	*2.0*	*2.3*	*2.1*	*1.8*	*d*
		JANZ	*2.6*	*3.6*	*1.5*	*1.2*	*d*
Wage rate w(t)	g_w	**USA**	*7.0*	*4.7*	*4.1*	*4.1*	*d*
		Canada	*8.2*	*4.7*	*3.4*	*2.8*	*d*
		Europe	*9.2*	*7.7*	*11.8*	*9.8*	*d*
		JANZ	*4.5*	*3.7*	*2.7*	*0.6*	*d*
Population L(t)	g_L	**USA**	*1.03*	*1.05*	*1.08*	*1.06*	*c*
		Canada	*1.06*	*1.40*	*1.13*	*0.93*	*c*
		Europe	*0.38*	*0.41*	*0.20*	*0.00*	*c*
		JANZ	*0.90*	*0.67*	*0.59*	*0.47*	*c*

a- Bureau of Economic analysis (*http://www.bea.gov/national/nipaweb/SelectTable.asp?Popular=Y*)
b- Bureau of Labor Statistics (*http://data.bls.gov/cgi-bin/surveymost?pr*)
c- UN Population Division (*http://esa.un.org/unup/*)
d- OECD statistics (*http://stats.oecd.org/Index.aspx*)

Figure C-1: *Urban population for the metro-regions of USA, Europe and OECD Pacific over 1980-2001*[1]



[1] For the sake of legibility of the graphs, we report results only the ten largest agglomerations in the USA and Europe.

IV- Long term equilibrium

Figure C-2: *Dynamic trends of the urban model of USA over 1000 periods for* $\phi = 1$[2]

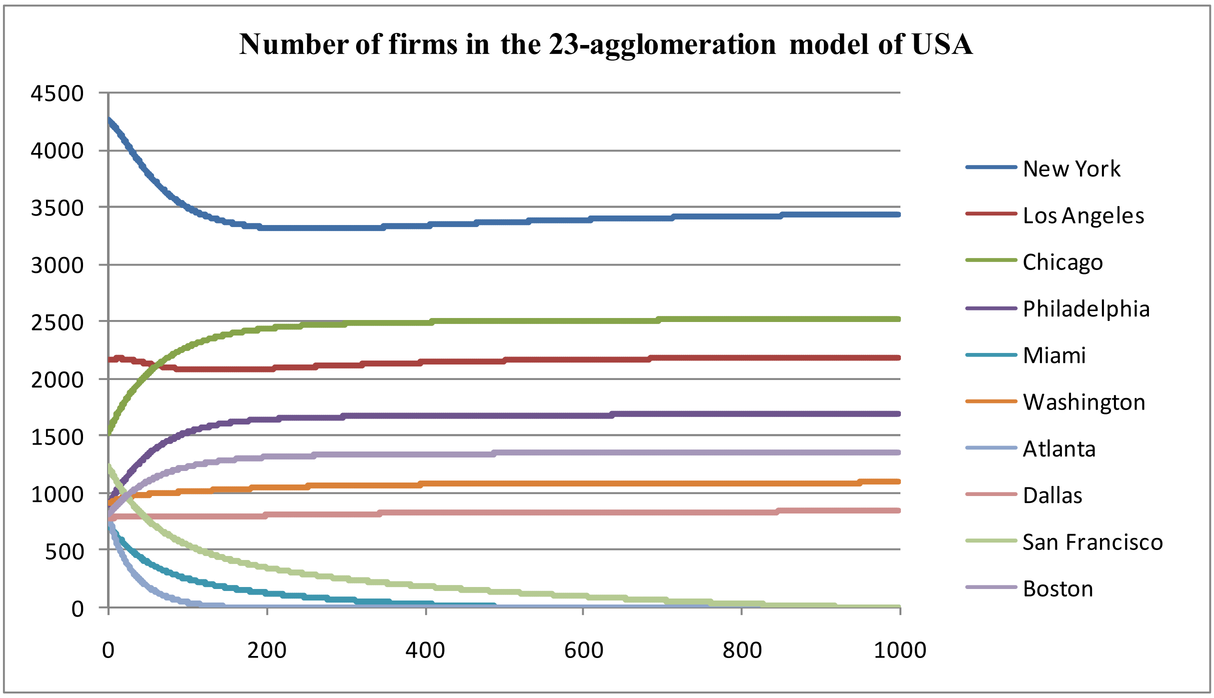
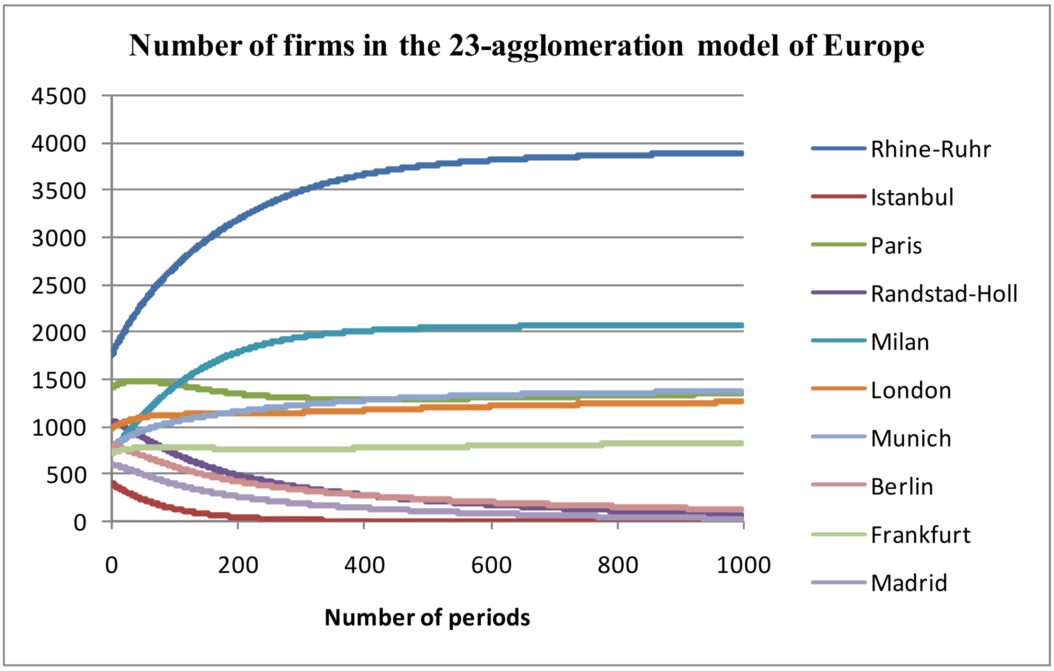


Figure C-3: *Dynamic trends of the urban model of Europe over 1000 periods for* $\phi = 1$[3]



[2] For the sake of legibility of the graphs, we report results only the ten largest agglomerations in the USA and Europe.
[3] For the sake of legibility of the graphs, we report results only the ten largest agglomerations in the USA and Europe.

Figure C-4: *Dynamic trends of the urban model of OECD Pacific over 1000 periods for* $\phi = 1$

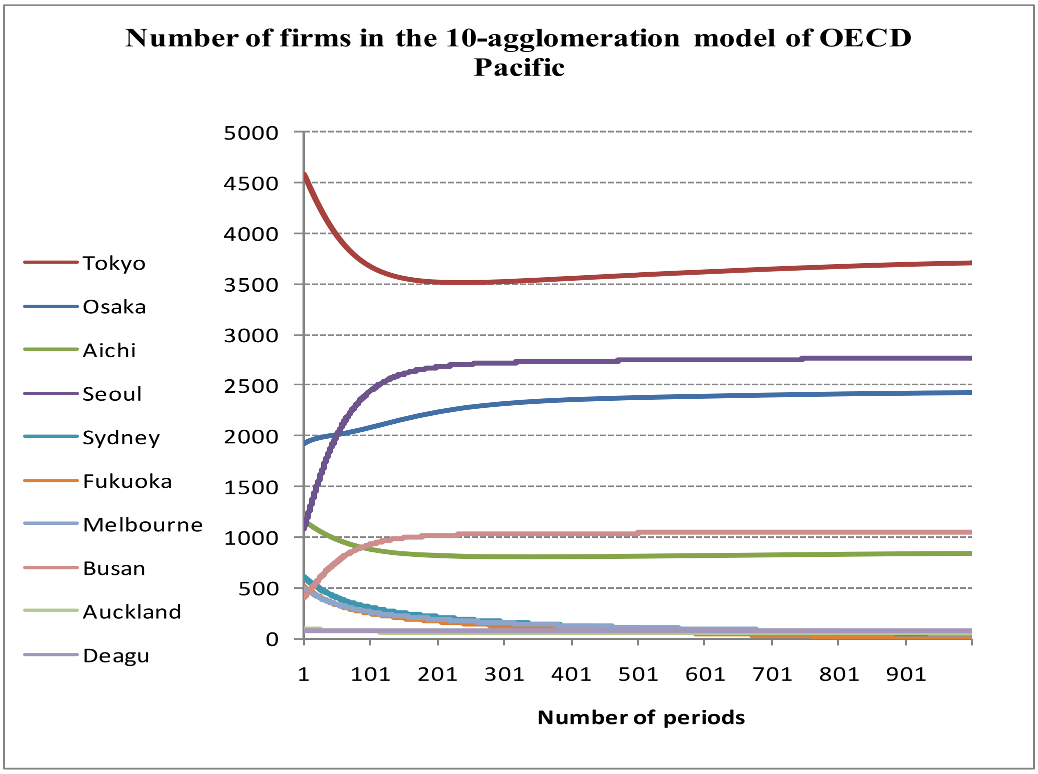


Figure C-5: *Bifurcation diagram for USA*

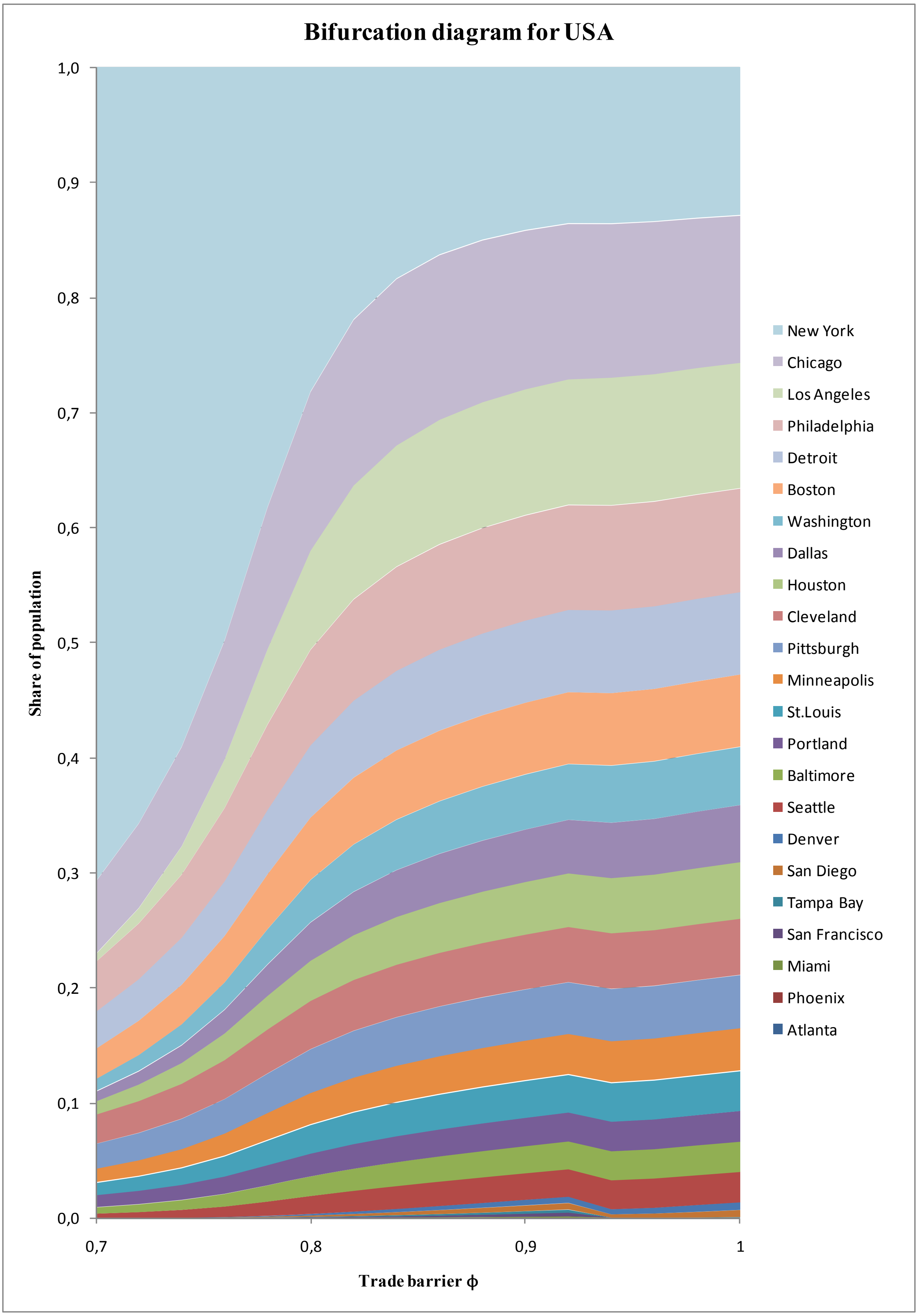


Figure C-6: *Bifurcation diagram for Europe*

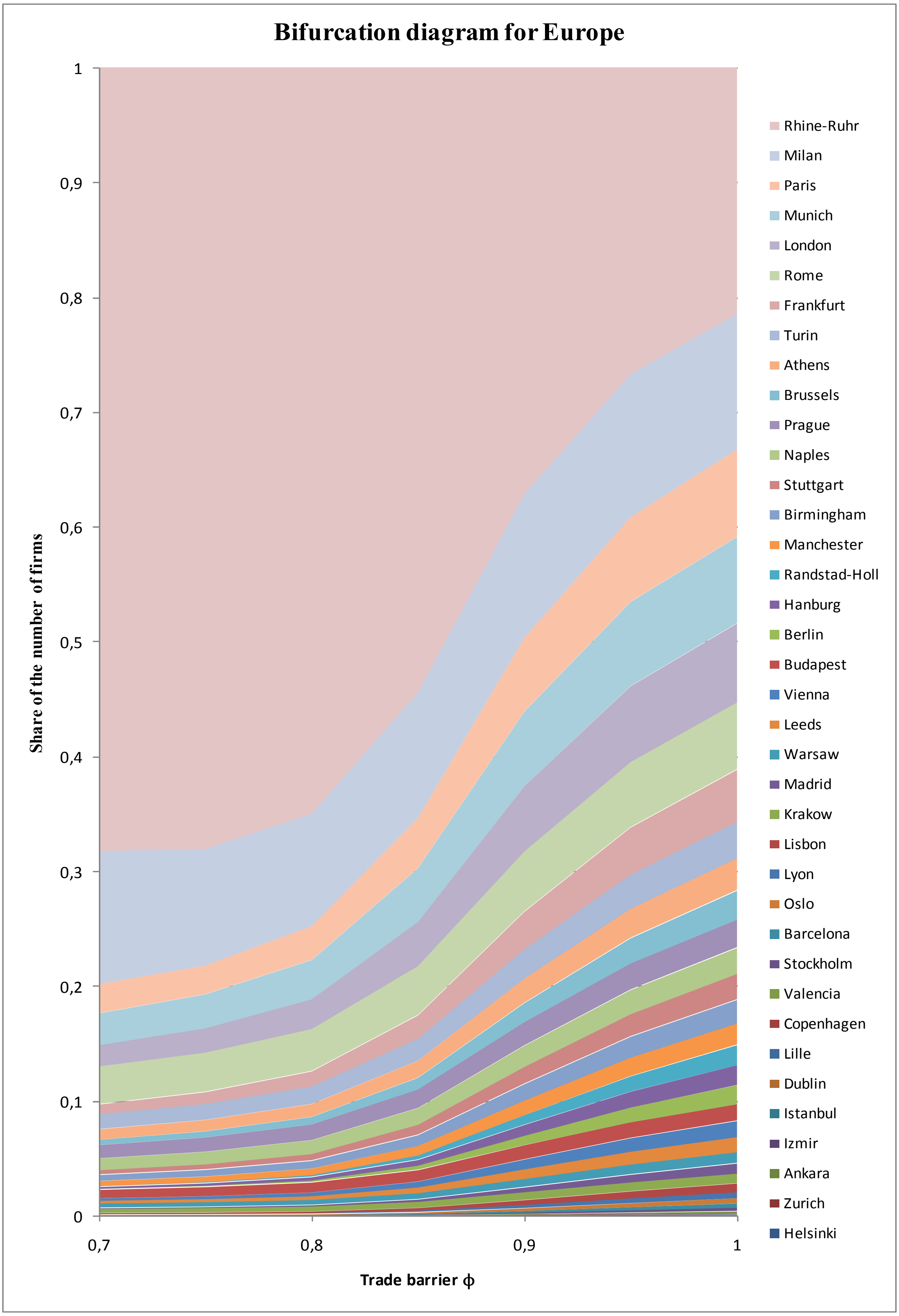
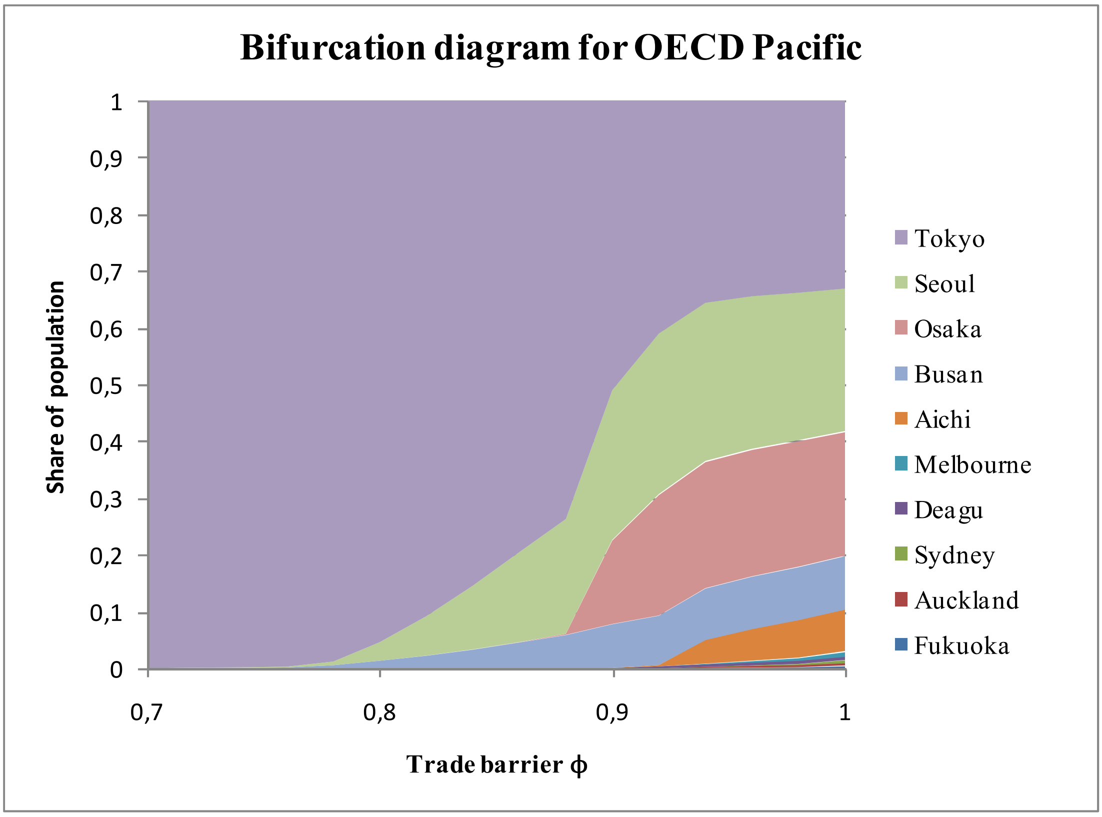


Figure C-7: *Bifurcation diagram for OECD Pacific*



V- Forecasting exercice

Table C-4: *Average growth rate of macroeconomic variables over the period 2001-2050.*

Macroeconomic aggregates		Region	2001	2010	2020	2030	2040	2050
Production value	*10^{12}\$*	*USA*	18.2	22.6	26.8	31.6	36.8	43.6
		Canada	1.3	2.0	2.5	3.3	3.9	4.6
		Europe	17.1	22.9	28.0	34.9	37.6	40.6
		OECD Pacific	9.1	12.2	14.9	17.6	18.6	19.4
Effective labor	*Million workers*	*USA*	142.6	161.9	165.7	161.5	161.9	165.6
		Canada	16.0	16.8	19.5	20.3	20.7	21.1
		Europe	299.4	287.4	294.8	293.7	282.3	267.3
		OECD Pacific	105.5	100.1	104.1	100.3	93.7	85.8
Average wage rate	*10^{3}\$/y*	*USA*	44.5	53.1	59.1	67.2	79.0	95.3
		Canada	23.3	29.8	43.7	61.5	75.0	88.5
		Europe	14.8	17.7	23.4	32.8	37.9	42.4
		OECD Pacific	25.0	31.0	44.7	60.1	69.3	78.7
$\overline{Pop(t)}$	*millions*	*USA*	285.3	304.6	328.2	344.2	351.0	355.1
		Canada	31.1	33.0	35.2	36.6	37.2	37.5
		Europe	588.2	596.6	600.2	597.6	587.8	574.0
		OECD Pacific	204.7	208.7	208.2	204.0	198.6	190.2

Table C-5: *Population forecasts for the 74 metro-regions over the period 2001-2050.*

	POPULATION (million)				AAGR (%)				POPULATION (million)				AAGR (%)		
	2001	2010	2025	2050	00/10	10/25	25/50		2001	2010	2025	2050	00/10	10/25	25/50
EUROPE								**USA**							
Rhine-Ruhr	13.4	14.3	15.5	13.1	0.7	0.6	-0.7	New York	18.5	21.5	20.5	18.4	1.7	-0.4	-0.4
Paris	11.1	11.3	11.6	9.9	0.2	0.2	-0.7	Los Angeles	12.6	15.1	14.9	14.1	2.1	-0.1	-0.2
Istanbul	11.6	10.7	9.1	5.5	-0.9	-1.2	-2.1	Chicago	9.2	11.9	13.1	14.2	2.9	0.7	0.3
Rand-Holland	7.4	7.3	6.9	4.7	-0.2	-0.4	-1.5	Philadelphia	5.7	7.6	8.6	9.8	3.2	0.9	0.5
London	7.3	7.5	7.9	6.8	0.3	0.3	-0.6	Miami	5.1	5.5	4.5	3.3	0.8	-1.4	-1.4
Milan	7.4	8.0	9.3	9.9	0.9	1.1	0.3	Washington	4.9	6.2	6.4	6.4	2.6	0.3	0.0
Berlin	6.0	6.0	5.6	3.5	0.0	-0.5	-1.9	Atlanta	4.4	4.1	2.6	1.2	-0.8	-3.1	-3.1
Munich	6.0	6.3	6.8	5.9	0.5	0.5	-0.5	Dallas	5.3	6.8	7.0	6.9	2.7	0.2	-0.1
Madrid	5.2	5.1	4.8	3.4	-0.2	-0.4	-1.4	San Francisco	4.2	4.5	3.8	2.8	0.7	-1.2	-1.2
Frankfurt	5.5	5.7	5.8	4.7	0.3	0.2	-0.9	Boston	4.4	5.8	6.5	7.1	3.1	0.8	0.4
Barcelona	4.7	4.6	4.2	2.7	-0.4	-0.7	-1.8	Houston	4.8	6.2	6.5	6.5	2.8	0.3	0.0
Hanburg	4.6	4.6	4.5	3.3	0.0	-0.2	-1.3	Detroit	4.5	5.9	6.6	7.5	3.0	0.9	0.5
Athens	3.9	4.1	4.5	4.3	0.5	0.6	-0.1	Phoenix	3.4	3.3	2.2	1.2	-0.4	-2.7	-2.6
Rome	3.7	4.1	5.2	6.7	1.2	1.6	1.1	Minneapolis	3.0	3.9	4.2	4.4	2.9	0.5	0.2
Brussels	3.7	3.8	3.9	3.2	0.2	0.2	-0.8	Seattle	3.1	3.8	3.9	3.7	2.4	0.0	-0.2
Ankara	4.0	3.6	2.9	1.6	-1.2	-1.6	-2.5	San Diego	2.9	3.2	2.8	2.3	1.3	-0.9	-0.9
Izmir	3.4	3.0	2.3	1.2	-1.5	-1.9	-2.8	St.Louis	2.7	3.5	3.7	3.8	2.6	0.4	0.2
Zurich	2.2	1.8	1.1	0.3	-2.4	-3.4	-5.1	Baltimore	2.6	3.2	3.2	3.2	2.3	0.1	-0.1
Lisbon	2.7	2.7	2.6	2.0	0.0	-0.2	-1.2	Denver	2.2	2.5	2.2	1.8	1.4	-0.8	-0.8
Warsaw	3.0	3.1	3.6	4.6	0.5	1.0	1.0	Tampa Bay	2.4	2.7	2.3	1.8	1.1	-1.1	-1.1
Copenhagen	2.4	2.2	1.9	1.0	-0.6	-1.2	-2.6	Pittsburgh	2.4	3.3	3.9	4.7	3.5	1.3	0.8
Budapest	2.8	3.2	4.2	7.3	1.3	2.1	2.3	Cleveland	2.1	3.1	4.0	5.2	4.3	1.8	1.1
Stuttgart	2.6	2.7	2.8	2.3	0.3	0.2	-0.8	Portland	2.0	2.7	3.0	3.3	3.5	0.9	0.4
Manchester	2.5	2.6	2.7	2.3	0.4	0.3	-0.7	**CANADA**							
Prague	2.3	2.6	3.2	4.1	1.2	1.6	1.1	Toronto	4.9	5.4	5.7	5.1	1.2	0.4	-0.5
Birmingham	2.6	2.7	2.8	2.4	0.4	0.4	-0.6	Montreal	3.5	4.2	5.1	5.7	2.0	1.4	0.5
Stockholm	2.1	2.1	1.9	1.3	-0.4	-0.6	-1.6	Vancouver	2.1	2.3	2.4	2.2	1.1	0.4	-0.4
Vienna	2.1	2.2	2.2	1.7	0.2	0.0	-0.9	**OECD PACIFIC**							
Naples	3.1	3.4	4.2	5.3	1.2	1.5	1.0	Tokyo	33.7	32.4	27.8	14.2	-0.4	-1.1	-2.8
Lille	2.6	2.4	2.1	1.2	-0.5	-1.0	-2.3	Seoul	22.1	24.2	31.9	37.5	1.0	2.0	0.7
Valencia	2.2	2.1	1.9	1.1	-0.4	-0.8	-2.0	Osaka	17.0	16.9	15.4	8.7	-0.1	-0.7	-2.3
Leeds	2.1	2.1	2.2	1.8	0.3	0.2	-0.8	Aichi	9.0	8.5	7.1	3.5	-0.6	-1.2	-2.9
Krakow	2.1	2.4	3.0	4.0	1.2	1.6	1.2	Busan	7.8	8.6	11.4	13.4	1.1	2.0	0.7
Turin	2.2	2.4	2.8	3.0	1.0	1.2	0.4	Fukuoka	5.0	4.6	3.5	1.5	-1.1	-1.8	-3.5
Helsinki	1.8	1.5	1.1	0.5	-1.5	-2.1	-3.4	Sydney	4.1	3.8	2.8	1.1	-1.0	-2.0	-3.8
Oslo	1.7	1.7	1.6	1.1	-0.2	-0.4	-1.4	Melbourne	3.5	3.2	2.4	1.0	-0.9	-1.9	-3.7
Dublin	1.5	1.4	1.2	0.7	-0.6	-1.1	-2.4	Deagu	2.5	2.4	2.5	2.1	-0.6	0.3	-0.6
Lyon	1.6	1.6	1.6	1.2	0.0	-0.1	-1.1	Auckland	1.2	1.1	1.0	0.7	-0.8	-0.6	-1.8

Table C-6: *Production forecasts for the 74 metro-regions over the period 2001-2050.*

Share of production in each agglomeration for the four OECD regions (%)					
Agglomeration	2001	2050	Agglomeration	2001	2050
EUROPE			USA		
Rhine-Ruhr	10.1	12.0	New York	20.3	17.4
Paris	8.6	9.4	Los Angeles	10.6	10.3
Istanbul	2.4	1.4	Chicago	7.8	10.3
Randstad-Holland	6.2	4.8	Philadelphia	4.6	6.8
London	5.9	6.8	Miami	3.6	2.0
Milan	4.6	7.5	Washington	4.8	5.4
Berlin	4.5	3.2	Atlanta	4.1	1.0
Munich	4.7	5.6	Dallas	4.2	4.7
Madrid	3.6	2.9	San Francisco	5.7	3.3
Frankfurt	4.3	4.4	Boston	4.2	5.8
Barcelona	3.2	2.2	Houston	3.9	4.5
Hanburg	3.5	3.0	Detroit	3.8	5.4
Athens	1.7	2.3	Phoenix	2.5	0.8
Rome	1.7	3.8	Minneapolis	2.6	3.2
Brussels	2.7	2.8	Seattle	2.8	2.8
Ankara	0.7	0.3	San Diego	2.2	1.5
Izmir	0.7	0.3	St.Louis	2.0	2.4
Zurich	3.7	0.6	Baltimore	2.0	2.1
Lisbon	1.5	1.4	Denver	2.2	1.6
Warsaw	0.6	1.2	Tampa Bay	1.6	1.0
Copenhagen	2.7	1.4	Pittsburgh	1.7	2.9
Budapest	0.4	1.4	Cleveland	1.5	3.2
Stuttgart	2.1	2.2	Portland	1.2	1.8
Manchester	1.4	1.6	CANADA		
Prague	0.7	1.6	Toronto	52.5	45.1
Birmingham	1.5	1.7	Montreal	29.3	39.3
Stockholm	1.9	1.4	Vancouver	18.1	15.6
Vienna	1.8	1.8	OECD PACIFIC		
Naples	0.7	1.5	Tokyo	41.0	28.0
Lille	1.7	1.0	Seoul	11.3	31.1
Valencia	1.4	0.9	Osaka	17.3	14.3
Leeds	1.2	1.3	Aichi	10.4	6.7
Krakow	0.3	0.6	Busan	4.1	11.5
Turin	1.1	1.9	Fukuoka	4.5	2.2
Helsinki	1.8	0.6	Sydney	5.3	2.4
Oslo	1.7	1.4	Melbourne	4.3	2.0
Dublin	1.7	1.0	Deagu	0.8	1.1
Lyon	1.1	1.0	Auckland	0.9	0.8

TABLE OF CONTENTS